万卷方法 | 社会科学研究方法经典译丛

主　编　沈崇麟　夏传玲

第 2 版

SOCIAL NETWORK ANALYSIS

A HANDBOOK

社会网络分析法

■ 约翰·斯科特　著　■ 刘　军　译　■ 沈崇麟　校

重庆大学出版社

版贸渝核字(2006)第 14 号

图书在版编目(CIP)数据

社会网络分析法/(美)斯科特(Scott, J.)著;刘军译.—重庆:重庆大学出版社,2007.1
(万卷方法.社会科学研究方法经典译丛)
书名原文:Social Network Analysis: A Handbook
ISBN 978-7-5624-2147-4

Ⅰ.社… Ⅱ.①斯…②刘… Ⅲ.社会关系—研究 Ⅳ.C912.3

中国版本图书馆 CIP 数据核字(2007)第 003603 号

社会网络分析法(第 2 版)
约翰·斯科特 著
刘 军 译
沈崇麟 校
责任编辑:雷少波 陈 进　　版式设计:雷少波
责任校对:李定群　　责任印制:张 策

*

重庆大学出版社出版发行
出版人:张鸽盛
社址:重庆市沙坪坝正街 174 号重庆大学(A 区)内
邮编:400030
电话:(023) 65102378 65105781
传真:(023) 65103686 65105565
网址:http://www.cqup.com.cn
邮箱:fxk@cqup.com.cn(市场营销部)
全国新华书店经销
重庆升光电力印务有限公司印刷

*

开本:787×1092 1/16 印张:12.25 字数:213千
2007 年 1 月第 1 版 2007 年 1 月第 1 次印刷
印数:1—3 000
ISBN 978-7-5624-2147-4 定价:28.00 元

作译者简介

约翰·斯科特(John Scott)　英国埃塞克斯(Essex)大学社会学教授。出版30余部书，论文近130篇。主要出版物有《企业、阶级和资本主义》(*Corporations, Classes and Capitalism*, 1985),《谁支配英国》(*Who Rules Britain?* 1991),《社会学理论》(*Sociological Theory*, 1995),《分层和权力》(*Stratification and Power*, 1996),《社会网络分析》(*Social Network Analysis*, 2000),主编《牛津社会学词典》(*Oxford Dictionary of Sociology*, 2005年第3版)。他是《社会学评论》(*Sociological Review*)杂志的编委,《英国社会学杂志》(*British Journal of Sociology*)和《欧洲社会理论杂志》(*European Journal of Social Theory*)的编委会成员。最近的著述有《社会理论:社会学中的核心问题》(*Social Theory: Central Issues in Sociology*,2006),主编《当代50位重要社会学家》(*Fifty Key Sociologists: The Contemporary Thinkers*,2006)。

刘军　博士后，教授，哈尔滨工程大学人文社会科学学院社会学系主任。2003年毕业于北京大学社会学系，获博士学位。北京大学2006年优秀博士学位论文二等奖获得者。主要著述:《社会网络分析导论》(2004)、《法村社会支持网络——一个整体研究的视角》(2006)、《社会网络模型及其哲学意义研究》(黑龙江大学博士后工作报告,2006年6月),发表"一般化互惠——测量、机制及方法论意含"(《社会学研究》2007年第1期)等论文近20篇。主要研究领域:社会网络分析、社会学方法论。完成国家社科基金项目、教育部人文社会科学研究项目、中国博士后科学基金项目各一项。目前主持教育部留学回国人员科研启动基金项目和2005年度黑龙江省留学回国人员择优资助项目。

总序 社会研究方法的现状及其发展趋势

近年来,社会调查技术和社会研究方法都有很大的发展。在调查技术方面,自上世纪70年代以来,社会变迁多次横断面的跟踪调查研究,几乎成为所有国家和地区了解社会结构转变和社会发展状况的基础性调查。这种调查不仅对社会学的研究有很大促进,对整个社会科学的研究都产生了重大影响,而且这些调查结果有的已成为政府有关部门决策的重要依据。国际上比较著名的此类调查有:美国芝加哥大学全国民意调查中心(National Opinion Research Center,简称NORC)的"社会综合调查(General Social Survey ,简称GSS)",英国埃塞克斯大学调查中心进行的"全国家庭生活和社会变迁调查",法国经济和社会调查所进行的"全国经济社会调查",日本社会学会组织进行"全国社会分层与社会流动调查(简称SSM)"。中国台湾"中央"研究院社会学研究所,也每两年进行一次"台湾社会变迁基本调查"。美国的"社会基础调查",现在已成为年度性的调查项目,它是美国国家基金会目前资助的最大的社会科学研究项目。以上这些调查,除美国的调查外,一般均因经费原因采用纵向的间隔性重复调查法,即每隔一段时间,进行一次全国规模的抽样调查。每次调查除保留社会研究所需的基本项目外,每次都有不同的主题。在间隔若干时间后,再重复同一主题的调查,这样的研究设计,使社会变迁研究在可以涉及更为广泛的研究领域的同时,具有更好的积累性和可比性。多年来,这些基础性调查获得的资料,滋养着大批的社会科学研究者,有时一项调查就有很多名博士生用来写博士论文,以此所取得的研究成就,其可靠性受到社会科学界的广泛认同。例如1997年出版的,以台湾社会变迁基本调查数据为基础的研究报告集《90年代的台湾社会,社会变迁基本调查研究系列二》收论文16篇,内容涉及社会生活的各个方面,在台湾引起了极大的反响。

国内社会科学界在这方面也有了长足的发展。笔者所在的中国社会科学院社会学研究所的社会调查和方法研究室,组织或参与了多项与社会变迁有关的大规模抽样调查,取得了一定的研究成果,并积累了大量有关社会变迁的宝贵数据资料,其中主要有:

1. 城乡家庭变迁系列调查:该课题是由中国社会科学院社会学研究所牵头,联合北京大学和地方社科院的研究人员展开的一项类似多次横断面的城乡家庭变迁调查。这一调查始于1981年的"中国五城市婚姻家庭调查",而后有

1988年的"中国农村家庭调查"、1991年的"中国七城市家庭调查"、1998年的"中国城乡家庭变迁调查"。

2. 有关中国城乡社会变迁的系列调查：调查始于1991年的第二批国情调查，然后有1992年的"中国城乡居民生活调查"、1993年的"第三批国情调查"、1995年的"第四批国情调查"和1997年的"中国沿海发达地区社会变迁调查"。上述调查虽然还不是严格意义上的多次横断面的纵贯研究，但研究者已在研究设计中尽量考虑到纵贯研究的基本原则，如调查队伍的稳定、指标的可比性和样本空间的延续性等。

3. 中国城乡社会变迁调查：这一调查开始于2000年，为中国社会科学院重大课题。目前已经完成第一期第一次调查和第二次调查，今后将把这一调查发展为连续的、定期进行的社会变迁调查。

在纵向调查技术取得长足进步的同时，上世纪末至今，电话调查技术也有很大发展。电话调查涉及的范围几乎与个别(面对面)访谈同样全面。电话调查中使用的一系列方法，是在20世纪70年代后期和面对面调查一起发展起来的。在20世纪80年代中，电话调查开始变得很普遍，且成为许多场合中各种调查方法的首选。正如某些学者所言，一种在公共和私营部门被人们用来帮助提高决策效率的收集信息的有效方法为人们所普遍认同时，这一现象本身就具有方法论上的意义。不仅如此，电话调查还有很大的实践意义，因为它为研究者提供了更多的控制调查质量的机会。这一机会包括抽样、被调查人的选择、问卷题项的提问、计算机辅助电话访谈(CATI)和数据录入。正因为如此，今天在各种社会调查中，如果没有发现其他重要的足以放弃使用电话调查的原因，电话调查由于其独特的对调查质量进行全面监控的优点，常常成为各种调查方式的首选。由笔者翻译，重庆大学出版社出版的《电话调查方法：抽样、选择和督导》一书，也于2005年面世。

无论是纵向调查抑或电话调查，实际上都是收集研究资料的方法，而应用社会科学的发展，不仅在于调查技术，即收集资料技术的发展，还在于研究方法和分析技术的发展。近年来，无论是定性研究方法，还是定量研究方法都有了长足的发展。

首先，计算机技术的发展可谓突飞猛进，它对当今社会生活的各个方面产生了巨大的影响，在悄悄地改变着社会科学的研究风格和研究方式的同时，也大大提升了社会科学学者的研究能力。这种影响表现在研究过程的各个阶段，从理论建构(概念映射)、问卷设计(专业的问卷设计软件)、调查实施(计算机辅助访谈、计算机辅助电话访问系统、网络在线调查系统)、数据录入(光学标记识别软件)到数据分析(包括文本、声音、图像资料的处理)，甚至延伸到写作发表阶段。这样的过程发生在如社会学、经济学、政治学、心理学、教育学中，促进了学科之间的相互借鉴和交叉融合，至少在研究方法上呈现这种趋势。随着计算机计算能力的大幅度提高，20世纪80年代之后，统计学领域内发生了一场"革命"，主要表现在对定类和定序变量的建模能力的大幅度提高上，以及与分

1988 年的"中国农村家庭调查"、1991 年的"中国七城市家庭调查"、1998 年的"中国城乡家庭变迁调查"。

2. 有关中国城乡社会变迁的系列调查：调查始于 1991 年的第二批国情调查，然后有 1992 年的"中国城乡居民生活调查"、1993 年的"第三批国情调查"、1995 年的"第四批国情调查"和 1997 年的"中国沿海发达地区社会变迁调查"。上述调查虽然还不是严格意义上的多次横断面的纵贯研究，但研究者已在研究设计中尽量考虑到纵贯研究的基本原则，如调查队伍的稳定、指标的可比性和样本空间的延续性等。

3. 中国城乡社会变迁调查：这一调查开始于 2000 年，为中国社会科学院重大课题。目前已经完成第一期第一次调查和第二次调查，今后将把这一调查发展为连续的、定期进行的社会变迁调查。

在纵向调查技术取得长足进步的同时，上世纪末至今，电话调查技术也有很大发展。电话调查涉及的范围几乎与个别（面对面）访谈同样全面。电话调查中使用的一系列方法，是在 20 世纪 70 年代后期和面对面调查一起发展起来的。在 20 世纪 80 年代中，电话调查开始变得很普遍，且成为许多场合中各种调查方法的首选。正如某些学者所言，一种在公共和私营部门被人们用来帮助提高决策效率的收集信息的有效方法为人们所普遍认同时，这一现象本身就具有方法论上的意义。不仅如此，电话调查还有很大的实践意义，因为它为研究者提供了更多的控制调查质量的机会。这一机会包括抽样、被调查人的选择、问卷题项的提问、计算机辅助电话访谈（CATI）和数据录入。正因为如此，今天在各种社会调查中，如果没有发现其他重要的足以放弃使用电话调查的原因，电话调查由于其独特的对调查质量进行全面监控的优点，常常成为各种调查方式的首选。由笔者翻译，重庆大学出版社出版的《电话调查方法：抽样、选择和督导》一书，也于 2005 年面世。

无论是纵向调查抑或电话调查，实际上都是收集研究资料的方法，而应用社会科学的发展，不仅在于调查技术，即收集资料技术的发展，还在于研究方法和分析技术的发展。近年来，无论是定性研究方法，还是定量研究方法都有了长足的发展。

首先，计算机技术的发展可谓突飞猛进，它对当今社会生活的各个方面产生了巨大的影响，在悄悄地改变着社会科学的研究风格和研究方式的同时，也大大提升了社会科学学者的研究能力。这种影响表现在研究过程的各个阶段，从理论建构（概念映射）、问卷设计（专业的问卷设计软件）、调查实施（计算机辅助访谈、计算机辅助电话访问系统、网络在线调查系统）、数据录入（光学标记识别软件）到数据分析（包括文本、声音、图像资料的处理），甚至延伸到写作发表阶段。这样的过程发生在如社会学、经济学、政治学、心理学、教育学中，促进了学科之间的相互借鉴和交叉融合，至少在研究方法上呈现这种趋势。随着计算机计算能力的大幅度提高，20 世纪 80 年代之后，统计学领域内发生了一场"革命"，主要表现在对定类和定序变量的建模能力的大幅度提高上，以及与分

布无关的统计分析模型的发展之上，特别是基于“Resampling”（包括 Bootstrap、Jackknife、Monte Carlo 模拟等）的建模技术*。同时，计算能力的提高还带动了基于神经网络、动态模拟、人工智能、生态进化等新兴的分析和预测模型的发展。这些进展都为定量社会科学研究提供了更多的可供选择的工具。

亚德瑞安·E·拉夫特里（Adrian E. Raftery）依据社会学家所处理的数据类型，将定量社会学在美国的发展划分为三个时代：第一代起始于上世纪 40 年代，交互表是其主要处理对象，研究重点是关联度和对数线性模型；第二代起始于上世纪 60 年代，主要处理单层次的调查数据，Lisrel 类型的因果模型和事件史分析是其研究重点；第三代起始于上世纪 80 年代后期，开始处理诸如文本、空间、社会网络等非传统的数据类型，目前尚没有形成成熟的形态。拉夫特里的综述，虽然更强调定量社会学研究对统计学的贡献，但也大致勾勒出定量社会学在国外的发展脉络。

从分析模型的角度来看，定量分析在以下几个方向有了突破性发展：

1. 缺失值处理：由于社会生活的复杂性，社会调查数据常常出现缺失值，传统的处理方式是忽略这些缺失值，或者用均值替代。但现在则倾向于用多重插值法（multiple implation）或者其他基于模型的方法进行处理。这些技术的发展，不仅会增强我们对数据的处理能力，而且将改变我们设计问卷的方式。基于这些技术，我们在不增加被访者负担的前提下，大大增加调查问卷的内容：每个被访者只回答问卷的一部分，然后通过对缺失值的处理，获得他们对未回答部分的估值。

2. 非线性关系：线性假定是经典定量分析的一个常见假定，但在实际研究当中，线性假定只能被看作是对社会现实的一个逼近和简化。面对具体的研究数据，如果没有理论上的明确指引（不幸的是，我们常常没有中程理论的指引），我们是无法在线性模型和非线性模型之间做出取舍的。但 MARS 模型的出现，让我们可以从经验数据当中获得最为拟合的变量之间的函数关系，而不必预先做出线性假定。这样，理论思考和数据分析就可以实现一个互动的循环过程，定量分析就不单单是对理论和假设的简单证伪过程，而是理论思维一个重要组成部分。

3. 测量层次：20 世纪六七十年代的统计模型，大多要求数据的测量层次在定距以上，如因素分析，但社会学的调查数据却大多为定类或定序数据。对应分析、Loglinear、Logit、Logistic Regression、潜类分析、Ordinal Regression、Normal Ogive Regression 等统计模型的出现，大大提高了定量社会学处理定类和定序数据的能力。

4. 测量模型：基于文化、社会、心理和认知等方面的考虑，在社会学界仍有人对问卷调查在中国的效度提出质疑。抛弃“本土化”的文化执著，我们更应当

---

* 对于当前一些国内尚无公认译法的模型、软件等的名称，本“译丛”都只给出了英文的表述，以免造成称谓上的混乱。

关注的是问卷调查的项目反应理论(item response theory),即被访者回答问卷题器时的过程模型。这方面的进展主要表现在两个方面:一是分解测量量表的成分,如 Rasch model、IRT 分析、Mokken 分析等,二是将测量模型与因果模型或其他分析模型结合在一起,明确把测量误差引入到分析当中,充分评估它们对分析结果的影响,如结构方程模型。

5. 潜变量模型:与测量模型相关联的另外一个发展方向是潜变量模型,例如,潜变量分层分析(latent class analysis)、潜变量结构分析(latent structure analysis)、潜变量赋值分析(latent budget analysis)等。"潜变量"这一概念表明,我们可以通过测量"显变量"来测量无法直接观察的理论概念,如权力、声望、地位等。这样,理论和现实之间,通过"潜变量"到"显变量"的映射(测量过程),就有了联接的桥梁。

6. 分析单元的层序性:在定量分析当中,我们常常强调要避免出现"生态谬误",即分析单元的层次和结论或推论的层次不一致。与其相关的方法论争论是"宏观和微观"的问题。随着多层次模型的出现,我们可以同时考察多个层次上的问题,我们可以把个人放在其家庭背景中,再把家庭放在社区的背景下,考察个人层次的变量对社区变量的效应,或者社区层次的变量对个体行为的具体影响。在定量分析模型当中,"宏观和微观"的联接获得了建模技术上的支持。在这个领域当中,还有一个方向也值得关注:分析宏观层次的数据,对微观层次进行推论。

7. 社会网络模型:区分"关系数据"和"属性数据",是把分析重点从个体/群体等社会单元转移到这些社会单元之间关系的第一步,社会网络模型是目前发展较快的一个定量分析领域,其理论根基是结构主义。社会网络分析目前仍然具有较浓厚的"形态学"特征(基于图论的缘故),但却为我们理解社会关系在社会空间上的形态奠定了基础,通过计算机模拟和研究社会网络的历期数据,研究社会结构的"发生学"性质模型也处在萌芽状态当中。

8. 系统动力学:如果说社会网络模型是在社会空间上拓展定量社会学的研究手段,那么,社会过程在时间上和物理空间上的属性,则是事件史模型、事件数模型、历期分析、Cox 回归、时间序列分析、Cohort 分析、状态空间模型等模型的研究对象。在这个领域,计量经济学为定量社会学研究提供了许多有益的范例。

9. 预测模型:上述模型仍然是在分析主义的范式下。有些社会学的应用研究,更强调模型的预测精度,而不是模型的认知价值,例如,社会趋势的预测。由于计算能力的提高,神经网络、基因算法、人工智能、模式识别等数据挖掘技术有了长足发展,已经出现了许多拟合经验数据的预测模型,比较成功的应用出现在计量经济学领域(如对股市的预测)。

10. 计算机模拟:对于社会学应用研究而言,研究的对象具有历史性、规模大、变迁的过程不仅漫长且表现某种渐进性,且因社会隔离/社会伦理原因无法接近或有实验禁忌等,无法直接进行观察和研究,这时计算机模拟就成为一个可供选择的替代方案。计算机模拟主要有两个类型,一是基于计算机网络的模

拟:每台微机作为一个代理,整个网络作为“社会”,实时演化,如法国的Swarm计划;二是基于概念模型的系统,在计算机时间上,按照既定规则运行,较有名的研究是罗马俱乐部的《增长的极限》,常见的软件有Simul,Arena等。自然科学家对此方向似乎比社会学家更有兴趣。

定性研究方法一直是社会学研究领域中比较传统的研究方法,在社会学研究的古典时期,它甚至是社会学家手中唯一的研究方法。但随着定量研究方法在社会学研究中的广泛应用,定性研究方法就似乎越来越不受到人们的重视。但需要澄清的事实是,在定量分析模型取得飞速发展的同时,在过去的二十多年里,定性研究方法也有了长足的进步。主要表现在以下六个方面:

1.研究素材日益扩大:除了传统的参与观察、深度访谈、专题小组访谈之外,会话、交谈、电视、广播、文档、日记、叙事、自传(autobiography)等社会过程中自然产生的素材,甚至社会学理论本身(理论的形式化),也开始进入定性分析的视野当中。所有这些资料,不仅可以以文本的格式存储,而且,新型的多媒体介质,如图像、声音和视频,作为原始的分析素材,也日益成为定性分析的新宠。

2.分析方法更加多样:定性方法的种类在最近的二十多年中,更是有了一个质的飞跃。在比较传统的、源自语言学的方法,如内容分析、话语分析、修辞分析、语意分析、符号学、论据分析等方法之外,社会学家也创造出自己独特的定性分析方法,如施特劳斯(Strauss)等人的扎根理论、海斯(Heise)的事件结构分析、拉津(Ragin)的定性对比分析、Abbott和Hrycak采用最优匹配技术的序列分析、亚贝儿(Abell)的形式叙事分析(formal narrative analysis)、鲍尔(Bauer)等人的语库建设、Attride－Stirling等人的主题网络分析和神经网络技术应用的定性分析领域。所有这些方法的一个共同特征是,把定性研究向更加系统、更加精确、更加严格、更加形式化的方向推进。

3.认识论基础更加多元化:现象学、释义学和本土方法论(ethnomethodology)的认识论,一直是定性分析的大本营,但近年来,实证主义也开始逐渐为定性分析所接纳,解释和阐释之间,由激烈的对立关系,逐渐演变为相互融合。

4.研究过程更加客观规范:定性分析的一个主要问题在于阐释过程中不可避免的主观性。为了尽可能消除“解释者偏见”和主观选择性,定性分析开始遵循严格的程序模板或程序规则,并尝试引入定量分析中的“信度”、“效度”、“代表性”等概念,通过编码和对比,再加上传统的定性分析标准,如可解释性、透明性和一致性,使得定性研究的过程更加规范、阐释的结果更加客观,研究的结论更加可信。

5.研究过程更加有效率:这主要应归功于大量计算机辅助定性数据分析(CAQDA)软件的涌现。从上个世纪80年代以来,定性分析过程的数字化和计算机化,已经是一个不可逆转的大趋势。这种发展趋势与定性研究者的理论取向无关,不管他们的理论立场是实证主义、符号互动论,还是本土方法论,大多数定性研究者都在自己的研究当中,开始采用计算机来辅助定性资料的分析过程。据不完全统计,目前已经有二十多种定性分析的软件,分别隶属于德国、英国、法国、美国等国家。其中,有一些软件是国外研究机构的科研成果,可以免费使用,但比较

成熟的定性辅助系统大多是商业软件。这些定性分析的辅助系统,不仅使得研究者从处理大量文字材料的繁复劳动中解放出来,而且能够让研究者共享他们各自分析的细节,从而改变了定性研究的流程和研究集体之间的合作方式。同时,由于采用数据库结构,定性资料的管理也更加方便,这就为组织大型定性研究项目(包括多个研究地点、多个研究对象、历时的定性研究)提供了新的可能性。越来越多的定性研究人员开始走出他们的摇椅,坐到计算机屏幕前,湮没在访谈资料和故纸堆中的定性社会学家的形象已经一去不复返了。

6. 定性研究和定量研究的结合更加紧密:在定量分析方法的教材中,定性研究常常被看作是定量研究的前期准备工作,但定性研究者却持完全相反的观点,他们一般认为定性方法是自成一体的,可以完成从形成概念到检验假设的全部研究过程。在实际的应用研究中,定性方法和定量方法常常是交织在一起的,例如,克劳(Currall)等人在研究组织环境重要的群体过程时,通过内容分析把5年的参与观察资料量化,然后用统计分析来检验理论假定。格雷(Gray)和邓斯坦(Densten)在研究企业的控制能力时,利用潜变量模型把定性方法和定量方法有机结合在一起。雅各布斯(Jacobs)等人在研究比利时的家庭形态对配偶的家庭劳动分工影响时,首先用定量方法对纵向调查数据进行分析,从定量分析的结果中,又延伸出对核心概念的定性研究。这三个研究分别代表了定量和定性方法相互融合的三个方向:①克劳等人的研究代表着定性方法的实践者试图将定性数据尽可能量化的取向,近年来涌现出的处理调查数据中开放题器的编码问题的工具软件(如 Words at,Smarttext 等,注意:它们都是由著名的统计软件公司出品的处理定性资料的软件),处理定性资料的传统内容分析软件(如 Nvivo、MaxQDA、Kwalitan 等)也开始提供将定性资料转换到常用统计软件的数据接口,这些工具上的革新将加快这种趋势的发展。②格雷和邓斯坦的工作代表了"方法论多元论"的取向,即在应用研究过程中,通过核心概念的测量模型,把定性研究和定量研究结合在一起。③雅各布斯等人的工作则代表了一部分定量研究者对过度形式化的定量方法的不满,并试图通过定性方法加以弥补。在定量研究领域中,对"模型设定"问题的关注,是定量方法重新试图返回定性研究这种取向的另外一种表现。

与社会调查技术和社会研究方法突飞猛进的现实相比,我国学术界在这些方面的论著的出版似乎显得有些迟缓。虽然已经翻译了美国的一小部分经典定量分析教材,如布莱洛克(Blalock)和巴比(Babie)的教材,也有自己编写的一些教材,如袁方等人的《社会研究原理和方法》、卢淑华的《社会统计学》等,此外,偏重软件操作的还有郭志刚的《社会统计分析方法——spss 软件应用》、郭志刚的《logistic 回归模型——方法与应用》、阮桂海的《spss for windows 高级应用教程》等。在《社会学研究》等专业杂志上,也常常有一些定量分析的应用研究,可是专门的方法和应用模型研究却没有,也没有专门的方法研究期刊。仅就定量研究方法的介绍而言,也存在一些缺陷,主要表现在:

1. 原理和操作脱节。

2. 过分依赖某些商业软件，不全面。

3. 与中国的实证研究相脱节。

4. 不能反映当前方法研究的最新进展。

与定量研究方法相比，由于各种原因，定性研究方法的引进和介绍都比较少。在福特基金会资助的方法高级研讨班上，曾讨论过一些定性研究方法。在定性方法研究方面也有少数专著，如袁方和王汉生1997年出版的教程，陈向明2000年出版的专著。但总体说来，我们对定性研究方法还停留在初步介绍的阶段，主要的介绍也局限在定性研究的研究设计和资料收集的阶段上，对定性分析方法的介绍则没有能够反映出当代定性方法的最新进展。特别是，在定性分析工具（定性分析软件）的引进和研究上，基本上还是一个空白。虽然不乏一些出色的定性研究报告，但从方法研究上讲，我们才刚刚起步。当然，我们同时还应当注意到，在历史学领域，我国对定性资料的鉴别、考据和分析，积累了大量的经验和知识，这也应当是定性方法研究的知识来源之一，应努力加以发扬光大。

令人欣慰的是，社会研究方法的引进和出版方面相对滞后的状况终于有所改观。重庆大学出版社的编辑，以独到的学术眼光，逆当前出版界唯利是图的不良选题风气，投入了大量的人力物力，组织出版"万卷方法"。自2004年至今，已引进社会科学研究方法方面的专著十余种，在我国社会科学界已经引起了一定的反响。然而，更为可贵的是，重庆大学出版社并未以已经取得的成绩而自满，而是再接再励，在原有"万卷方法"的基础上，进一步组织出版"万卷方法—社会科学研究方法经典译丛"。按我们的设想，"译丛"应该是一个开放的体系，旨在跟踪社会科学研究方法发展的前沿，引进和介绍这一方面的经典著作和最新成果。

"译丛"第一批有《抽样调查设计导论》、《社会科学研究设计原理》、《社会科学研究测量原理》、《社会科学研究分析技术》、《问卷设计手册》、《回归分析法》、《数据再分析法》、《分组比较的统计分析》、《社会网络分析法》、《广义潜变量模型》、《定性变量数据分析》和《复杂调查设计和分析方法》（书名也许有变化）等十余种，几乎囊括了研究设计、测量和分析方法的所有领域，涵盖从基础的回归分析到最前沿的潜变量分析和多水平模型等各种分析方法。无论是社会科学各专业的本科生、研究生，还是社会科学研究的学者都将从中有所收获。

"译丛"由中国社会科学院社会学所社会调查和方法研究室的多位研究人员担纲，主译者都是在社会研究方法各个领域中具有相当造诣的教师和研究人员。"译丛"的译者不仅仅把翻译看作是一个"翻译"，而且也把它看作是一次再学习和再创新。

我们期待"译丛"的出版能对社会研究方法的研究、应用和教学有所推动。

沈崇麟　夏传玲

2006年12月于中国社科院社会学所社会调查与方法研究室

译者前言

众所周知,中国社会是关系社会。如何研究关系? 视角当然多种多样,既可以像林语堂的小说中描述的那样对关系进行细致的刻画,又可以像黄光国等社会心理学家那样对人情、面子和关系网进行质的描述,更可以像在国外已经有七十多年发展历史的社会网络分析那样对关系进行量化的表征,从而揭示关系的结构,解释一定的社会现象。社会网络分析的意义在于,它可以对各种关系进行精确的量化分析,从而为某种中层理论的构建和实证命题的检验提供量化的工具,甚至可以建立"宏观和微观"之间的桥梁。

在社会网络分析领域,国内外已经出版了十余种导论性质的教材,本书可谓其中比较有代表性的一种。本书就像一本手册,引导读者进入社会网络分析的研究领域。它既适用于社会网络分析的初学者,也适用于对社会网络分析有所了解的人士。虽然本书第1章的最后一节很简要地介绍了本书概要,但是这里还有必要结合译者的体会补充介绍本书的内容,以便使读者对本书有一个更加深入的认识。

我们知道,常规统计学面对的数据都是属性数据,即社会行动者自身拥有的数据(如性别,年龄,GDP等)。用来分析属性数据的统计技术和方法(如回归分析、方差分析、对数线性模型等)也都比较成熟,但是这些方法大都要求自变量之间相互独立,否则会出现共线性问题,从而使得参数估计不准确。然而,当我们研究两个或多个行动者之间关系的时候,出现的则是关系数据,关系数据本身恰恰违反"独立性"这个假设,因而不能用常规统计学进行描述和推断(个体网的一些指标可利用常规统计方法来计算)。社会网络分析适用于分析关系数据,因而在很大程度上不同于常规的量化方法。因此,第1章明确地区分了关系数据和属性数据,从而为后续的章节奠定了数据分析的基础。

第2章比较详细地介绍了社会网络分析的发展,认为其起源于社会计量学、图论、数学、社会心理学、社会学和社会人类学的研究。最后指出社会网络研究主要分为整体网研究与局部网研究两大类。如何收集、整理关系数据? 这是第3章探讨的内容。本章探讨了关系数据的整理、存储和选择等问题。收集到的关系数据如何表征? 主要有两种方法,即图和矩阵。因此,第4章介绍了社群图和矩阵的有关知识,介绍了个体网络的密度和整体网密度的计算方法。

在描述网络特征的时候,既可以描述每个点的特点,又可以把该网络看成

是一个整体，进而描述整个网络的特征。怎样描述每个点的特征呢？一种重要角度就是计算与该点有直接联系的其他点有多少，这个值就是“中心度”；怎样描述作为一个整体的整个网络的特征呢？一种方法就是计算出一个指标，用它来描述该网络在多大程度上围绕着一个点而建立起来，这个值就是“中心势”。因此，第5章讨论了“中心度”和“中心势”指数，该章是对权力的量化研究。作者明确指出，中心度特指单个行动者的中心性，中心势特指一个图向核心点靠拢的程度，因而特指一个图的中心性。作者用一个社区结构研究的例子对中心性研究进行了展示。

第6章考察一些主要的子群概念，介绍了派系、成分、核、环和结群等子群概念。总的来讲，这些概念都是关于子群体的研究，主要用来揭示、解释一个网络中存在的子群体的情况。第7章从社会网络的角度对我们熟识的位置、角色和聚类等进行了分析。本章重点探讨了点的结构对等性、规则对等性、聚类、块模型等，并用实际例子进行了说明。在译者看来，社会网络分析领域关于角色的形式化研究是对社会学理论的重要贡献。第8章介绍一些用来展示关系资料的形式化研究。本章超出了简单的社群图描述，利用社会结构的多维“图景”来对网络进行可视化处理。最后，附录部分介绍了一些社会网络分析软件。由于到目前为止其中相当一部分网址都已经有所变动，因此译者尽可能查对并提供了新的地址，但是译者仍然不能保证这些网址就是最新的，因为随着时间的推移它们又可能有变动，请读者自行搜索。

本书既有社会网络理论上的分析，又有实际的例证和应用。本书作者尤其介绍了社会网络分析在亲属关系、社区结构、连锁董事会以及精英结构研究中的应用，展示了网络分析的应用潜力。相信读者可通过阅读这些内容体会到社会网络分析的魅力所在。

在译者看来，本书只是一本导读性质的书，这一点也是斯科特在本书开头明确表示过的。如果通过阅读本书，读者对社会网络产生了浓厚的兴趣，并且确实希望进行这方面研究的话，那么仅仅阅读本书就不够了。在如下几方面，需要读者体察：

首先，各种社会网络方法到底怎样应用？UCINET软件如何使用？读者不会从本书中得到过多的信息。从“即学即用”的角度上讲，尽管作者介绍了几种网络分析的软件，但是由于篇幅和本书的任务所限，作者不能对网络分析的具体操作技术进行细致的介绍。有关具体操作，请参见译者编著并即将出版的《整体网分析——UCINET软件应用》一书。

其次，社会网络分析近十几年又有重要进展。二方关系、三方关系、一般化的块模型、动态网络模型、p*模型、社会影响模型、社会选择模型、动态网分析等最新成果没有体现在本书当中。尽管这也不是本书的任务，但是对于希望进入网络分析前沿研究的学者来说，阅读本书之后，还要跟踪更新的文献，并有所创新。

再次，在译者看来，我们不应该仅仅把社会网络分析看成是一种工具或者

一套工具，而应该看成是一种方法论，即方法论的关系论。也可以把网络分析看成是一种关系论的思维方式。这种思维方式告诉我们，我们生活的世界是一个关系的世界，我们与他者是不可分的，是“共在”的，我们的思想、行为、生活不是独立的。这是一种关系论的视角，而不是简单的“他者”视角。这种视角恰恰体现了20世纪学术思潮的转向，即从实体论走向关系论。关系论的视角已经体现在哲学、政治学、经济学、社会学、心理学、人类学、国际关系学、管理学等很多学科中。

最后需要补充的是，如何解释社会网络研究的结果？我们认为，有时候仅仅从社会网络的角度给出的解释也是不够的。在数学领域中有一个著名的定理——哥德尔不完全性定理。该定理告诉我们，任何公设系统都不是完备的，其中必然存在着既不能被肯定也不能被否定的命题，任何完备系统都不能仅仅利用自身的定理来解释。在社会网络分析中又何尝不是如此呢！在一定意义上可以说，“网络只是管道”，管道的结构固然重要，但是管道内流通的内容也是我们不能忽视的。因此，社会网研究的结果需要结合抽象的“文化”、“地方性知识”等来说明。只有如此，我们才可能理解、解释社会网络现象。

说明及致谢

本人在2001年就读过斯科特教授的这本书，一直有把它翻译过来的想法。承蒙重庆大学出版社和中国社会科学院社会学所沈崇麟研究员的信任，本人才有翻译此书的机会。在翻译过程中，斯科特教授发给我本书的电子文档，加快了翻译的进度，特向斯科特教授致谢。沈崇麟研究员高水平的审校工作使本书增色不少，也使译者受益很多。当然，译文不当之处完全由译者本人负责。我的学生姜元杰协助我完成了第8章的翻译，陈卓、阎翔宇等帮我找出译稿中的一些不妥之处，本人在此对她们表示感谢。

本书的原文虽然不难，但是翻译出来成为符合中国人阅读习惯的书，这还是有点难度的。翻译中出现的失真不可避免。海德格尔甚至认为，任何翻译都是不可能的。不过我相信，社会网络毕竟存在于自然世界、人文世界、生活世界的各个角落，社会网络分析的技术也具有一定的“普适性”。因此，把它翻译为多国语言并为人所理解的，这应该是可行的。

国内的社会网络研究还刚刚起步，有关术语的翻译没有达成共识，本人对一些术语的翻译谨代表我个人的意见，不当之处，错漏之处在所难免，请读者指正。

刘　军

2006年12月25日

目录

网络和关系 1

作为一种比较新的技术，社会网络分析的潜力正在引起人们越来越多的兴趣。不幸的是，这种潜力对许多研究者来说是可望而不可及的，因为他们发现很难抓住这一方法的高深技术语言和数学语言，对这些技术的讨论要用到这些语言。那些曾经试图运用社会网络分析技术的人都是一些实践领域的研究者，拥有某些具体的研究兴趣，而总的来说，介绍这些技术的文献都是由具有数学功底的高级计算专家撰写的。很难找到一个恰当的计算机程序来进行社会网络分析，而一旦获得了程序，研究者又发现自己在如何运用该程序方面缺乏实践的指导。

本书的目的是试图填平理论和实践之间的鸿沟。我并不是一位接受过数学训练的专家，而是一位运用社会网络分析的研究者，因为在我主持的一项企业权力研究计划中特别需要进行网络数据分析。许多年来，我始终努力在一定程度上理解网络结构涉及哪些主要的测度，并且试图把数学翻译为简单的语言，也试图对特殊的模型与特定的研究需要之间的相关性进行评价。因此，本书的目的就在于根据这种经验对这些测度进行系统的总结，同时展示它们的应用。我不想把本书写成一本对社会学中的结构分析进行全面介绍的专著（参见 Berkowitz，1982），也不试图对现已发表的各种有关社会网络分析应用方面的文献进行评述（参见 Mizruchi and Schwartz，1987；Wellman and Berkowitz，1988）。我关注在评价社会结构之时所使用的一些关键概念——如密度（density）、中心度（centrality）、派系（cliques）等——并且试图把对这些概念的数学论述转译成比较容易理解的术语。

研究者必须理解他们使用的一些概念，这一点至关重要。例如，“派系”以及与之相关的一些概念都有众多不同的定义，所以研究者绝不能不

加分析地使用"唾手可得"的程序,不应假设该程序中的派系概念与研究者头脑中的概念是一致的。正是出于这个原因,我才在很多地方强调,测度的选择及其在特定研究主题中的运用恰恰需要践行中的研究者作出有根据的判断。这些理论问题和经验问题是不能回避的,不能仅仅根据得到部分理解的一些数学测度来回避这些问题。只有当研究者对某种特定测度的逻辑有清醒理解的时候,他(她)才能对该测度与某项特定的研究之间的关联做出一种有根据的社会学判断。

关系和属性

首先,我们必须定义最适用于进行社会网络分析的数据的类型。毫无疑问,那些对社会网络分析的应用感兴趣的人已经对此有一些想法:可用它来研究亲属模式、社区结构、连锁董事(interlocking directorships)等。但是关键在于,此类数据的共同特点可以得到明确的理解。我的观点是,社会网络分析适用于分析"关系数据(relational data)",而用来分析其他类型数据的那些技术对生成此类数据的研究来说,其价值似乎有限。

社会科学数据的最一般特点是,它们根植于文化价值和符号体系之中。与自然科学的物理数据不同,社会科学数据是通过意义、动机、定义和类型化(typification)建构起来的。众所周知,这意味着社会科学数据的生产涉及一个解释过程。在这种解释过程的基础上,社会科学家就形成了各不相同的数据类型,适用于分析这些类型数据的方法也各不相同。

数据主要分为"属性数据"和"关系数据"两类[1]。**属性数据**(attribute data)指涉及能动者(agents)的态度、观点和行为方面的数据,它们被视作为能动者的个人或者群体所具有的财产、性质、特点等。例如,通过调查和访谈收集到的资料常常被简单地看成是特定个体的一些属性,并且可以利用许多现有的统计程序对这些数据进行量化分析。适用于分析属性数据的方法主要是"**变量分析法**(variable analysis)",这些方法把各种属性测量为一些特定变量(收入、职业、教育程度等)的取值。

另一方面,**关系数据**(relational data)则是关于接触、联络、关联、群体依附和聚会等方面的数据,这类数据把一个能动者与另外一个能动者联系在一起,因而不能还原为单个行动者本身的属性。关系不是行动者的属性,而是行动者系统的属性;这些关系把多对行动者联系成一个更大的关系系统。适用于分析关系数据的方法就是**网络分析**(network analysis)。在网络

分析中，关系被认为是表达了能动者之间的关联。当然，尽管对这些关系可以进行定量的统计计量，网络分析也可以由一系列有关网络结构的定性测度构成。

尽管属性数据和关系数据在有关方法的文献中得到了充分的讨论，但它们并不是社会科学数据中仅有的两类。第三种类型是**观念数据**（ideational data），它描述的是意义、动机、定义以及类型化本身。尽管观念数据居于社会科学的核心地位，但是分析这类数据的技术没有分析前两类数据的技术那样完善。韦伯（Weber，1920-21）概括提出的**类型分析**（typological analysis）法是最富有成效的探讨，但是需要进一步发展（参见 Layder，1992）[2]。

尽管存在着各种不同的数据类型（如图 1.1 所示），每种数据各有其适当的分析方法，但是收集各种数据的方法并没有什么独特之处。例如，收集属性数据的方法和收集关系数据的方法之间没什么差别。这三类数据常常一起收集，成为同一项研究的各个内在部分。例如，有关政治态度的研究可能把态度与群体成员和共同体依附感二者联系在一起；或者对连锁董事会成员的研究可能把成员与各个公司的规模和获利联系在一起。在任一情况下，可以认为，问卷法、访谈法、参与观察法或者文献分析法等都可用于收集数据。

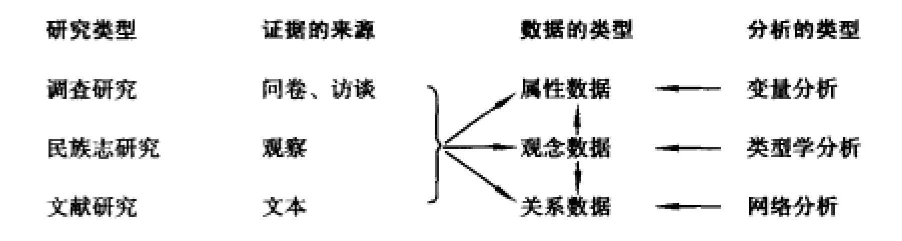


图 1.1 三类数据及其分析

例如，对友谊的研究倾向于遵循莫雷诺（Moreno，1934）的引导，即应用问卷法考察回答者对朋友的选择情况。在此类研究中，研究者仅仅请回答者指出他们的朋友，询问诸如“请说出你的 4 个最亲近朋友的名字”。当然，这种研究会引出一些方法论问题。人们发现，回答者难以回答一个选项不受限制的问题。有人认为自己说不出 4 个朋友的名字，一些人发现这种开放式的问题既耗时又乏味[3]。所以人们改用名册选择朋友的方法，问被调查者：“你认为下面哪些人是你的朋友？”这种研究要求研究者具有充分的知识，进行认真的准备，因为他必须对展示在被调查者面前的名册进行编辑。但是，这类方法的优点是可作调整，即要求回答者对自己与朋友之间的交往程度进行分类，从而表明关系的“紧密性”和重要性。然而，在

这两种情况下，有关认识和被调查者的合作的方法论问题与在收集有关态度和观点的信息中存在的问题完全一样。

在社会学传统所关注的一些主要问题中，关系数据都居于核心，因为社会学传统强调对社会行动的**结构**进行研究。结构建立在关系的基础上，因此，可以通过对关系数据的收集和分析来进行结构社会学研究。令人感到背谬的是，大多数现有的有关研究方法和数据收集方法的文献都很少关注这类数据，而关注用来研究属性数据的变量分析。社会网络分析的形式的、数学的技术和方法恰恰针对的是关系数据，但是，这些方法的发展和对它们的讨论却是在主流研究方法之外进行的。尽管结构分析方法已经有很多突破，但是对很多希望应用该方法的人来说，网络分析仍然可望而不可及。

最初，社会网络分析是在伟大的人类学家拉德克利夫-布朗（Radcliffe-Brown）对结构的关切之中，以一种相对非技术化的形式出现的。从 1930 年代到 1970 年代，越来越多的社会人类学家和社会学家开始构建拉德克利夫-布朗所说的“社会结构”概念，在此过程中开始认真思考社会生活中的“织构（fabric）”和“网（web）”这些隐喻。正是从这些目的在于理解社会行动据以组织起来的“相互交织”、“互相关联”的纺织式（textile）隐喻之中，社会“网络”这个隐喻才走到前台，研究者开始探讨社会网络的“密度（density）”和“组构（texture）”。然而，从 1950 年代开始，一小群专家开始用一些比较形式化的语言来转译这些隐喻。从 1970 年代初开始，在此领域出现了大量的技术性和专业应用方面的文献。社会网络分析的一些关键概念从这些文献中脱颖而出，从此以后，这类技术正式进入数据分析的主流，并且得到了广泛的应用。

本书概要

本书的目标是成为一本社会网络分析导引或者手册，而不是需要坐下来一口气读完的教材。我已经把一些辅助性的知识点和复杂的技术语言限定在脚注之中，但是有一定数量的复杂公式必须保留在正文之中。我希望这已经达到最少了。我建议社会网络分析的初学者首先阅读第 2、3 章，然后略读本书的其余内容，再返回到一些难点部分。对于那些已经熟悉社会网络分析的读者来说，他们的阅读过程可以与之相反，首先浏览第 2、3 章，然后重点阅读第 4 ~ 8 章的内容。如果在一项特定的研究中需要运用某种技术的话，那么与该技术相关的章节就应该重点阅读。尽管后续章节

依赖于前文中提出的论证,但是,当试图应用某种特殊技术的时候,每一章都可以看成是一个参考源。

第2章讨论社会网络分析的发展,认为其起源于群体社会心理学,随后在对工厂和社区进行的社会学和社会人类学研究中发展起来。本章重点探讨本书中出现的一些理论观点,并且指出它们与从1970年代以来出现的越来越多的技术复杂性是如何关联的。社会网络分析后来的这些发展体现在1970年代早期出现的如下两种标识性研究中:李(Lee,1969)对堕胎主义者的研究和格拉诺维特(Granovetter,1974)对谋职的研究。在第3章中,我考查在界定社会网络的界限以及在选择有待研究的关系的时候出现的一些问题。谈论这些问题的目的是引入社会网络分析的某些必备工具,笔者着重介绍了矩阵和社群图,并把它们看成是对关系数据进行模型化处理的简单且直观的方法。

第4章介绍社会网络的一些基本构成要素。本章从社会计量学的基本观念出发,即把网络表达为由"点"和"线"构成的图,并且指出如何据此提出一些概念,如"距离"、"方向"和"密度"等。第5章讨论了点的"中心度(centrality)"和整体网的"中心势(centralization)",并在第4章的论证基础上,指出如何从局部的、"个体中心"测度转移到整体的、"社会中心"的测度。第6章考察在分析社会网络内部子群的时候所使用的一些主要概念——即把网络分裂开来的"派系(cliques)"和"社会圈(circles)"。第7章探讨那些通过社会关系界定的位置,转而考察"位置"的结构以及这些位置如何结合成为复杂的"拓扑"结构。第8章介绍了一些用来展示关系数据的形式化方法,本章将超出简单的社群图,走向社会结构的多维"图景"的生成。最后,附录部分介绍并比较了一些主要的社会网络分析软件。

第4~8章每一章的结论部分都考察了所讨论的测度在特定经验研究中的应用。我们所评论的研究包括在亲属、社区结构、连锁董事会以及精英结构中的研究。之所以展示来自于这些主要研究者的成果,目的在于一窥社会网络分析带来的潜力。

注 释

1 这一区分取自韦尔曼(Wellman,1980)与伯科威茨和海尔(Berkowitz and Heil,1980)的早期论述。

2 也可参见埃布尔(Abell,1986),他做了一项有趣的尝试,即将图论技术应用于对行动的序列模式的分析。就其本身来讲,尽管这并不是一个社会网络分析的例子,但是图论对于社会网络分析来说仍然是基础。

3 "选择提名4个朋友"这种情况尽管比较常见,但也很武断。不管在问题中具体指定选择多少个朋友,类似这种武断性的问题仍然存在。

2 社会网络分析的发展

当代社会网络分析的发展得益于多种多样的学科和学派，这些学派在社会网络分析的发展过程中相互影响，时而汇聚交融，时而分道扬镳。但是，不管怎样，社会网络分析的主流学派在这个复杂的发展进程中清晰地构建起来。这个学派中主要有如下三个传统：社会计量学者，他们通过研究小群体，在技术上从很多方面推进了图论方法的发展；1930 年代的哈佛学者，他们研究了人际关系的模式，提出了“派系”这个概念；曼彻斯特的人类学家，他们在前两种传统的基础上考察了部落和乡村的“社区”关系结构。这些传统最终于 1960 年代和 1970 年代在哈佛大学又一次汇聚在一起，当代社会网络分析正是在那个时代出炉于哈佛（参见图 2.1）。

在 1930 年代，受沃尔夫冈·科勒（Wolfgang Köhler）的“格式塔（Gestalt）”理论的影响，一群移民自德国的学者在美国开展认知心理学和社会心理学的研究。这一工作促成了有关社群图和“群体动力学”问题的研究大量出现。他们利用实验方法或者类似实验的案例研究法考察了群体结构，探究了信息和观念在群体中的流动。与此同时，哈佛大学的人类学家和社会学家致力于推进英国社会人类学家拉德克利夫-布朗的一些思想。他们在工厂和社区研究方面产生了很多重要成果，这些成果强调了社会系统中非正式关系、人际关系的重要性。在英国尤其是在曼彻斯特大学，为了推进拉德克利夫-布朗的研究，一些学者重点分析了冲突和矛盾，并且用这一思想来研究非洲的部落社会，稍后研究了英国乡镇。在早期传统的基础上，这些作者把数学和实质性的社会理论结合起来，并有重大推进。然而，直到 1960 年代，社会网络分析才在方法论的完善方面有最终的突破。在哈佛大学，哈里森·怀特（Harrison White）开始扩展对社会结构的数学基础的研究，把他在北美的同事的许多洞见锻造综合在一起，并由他

的学生发扬光大。随着这些学生在世界各地的众多院系供职，怀特的思想和英国研究者的工作结合在一起，成为一种复杂但又越来越具有凝聚性的社会网络分析框架。

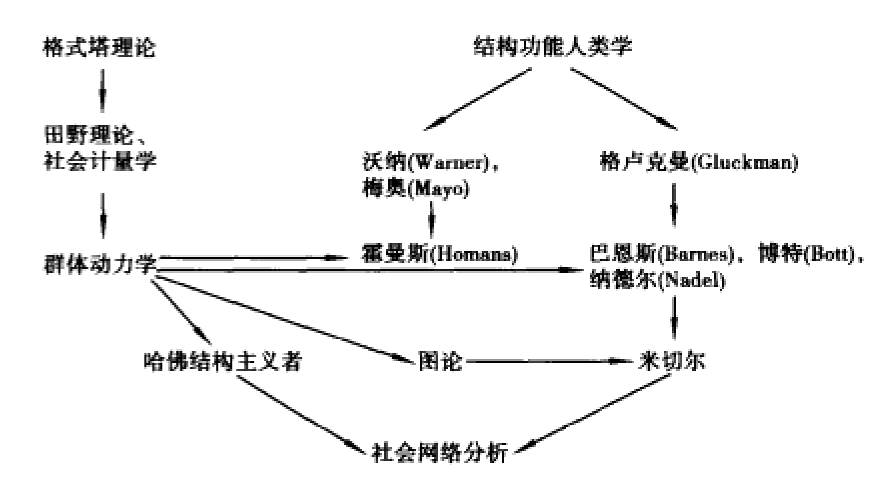


图 2.1 社会网络分析发展的系谱图

本章将勾勒出社会网络分析的三个传统，介绍哈佛大学哈里森·怀特小组的一些主导性的创新思想。这种评述也将凸现出在社会网络分析领域中持续存在的争论性课题，我将指出这些课题根植于社会学中的一些核心性、实质性问题。

社会计量分析和图论

心理学中的格式塔传统主要与科勒（参见 Köhler，1925）的工作有关，该传统强调一些组织化的模式，认为思想和观念是通过组织化模式才得以结构化的。他们把这些模式看成是“整体”或者系统，它们拥有不同于其“部分”，并进而决定“部分”之本质的独特性质。例如，人们往往根据一些特定的方式来理解个体对象，因为从字面上讲，这些个体先入为主地存在于人脑中复杂的、有组织的概念图式之中。对世界上各种事物的理解不能脱离这些心智图式，从根本上说，是由这些图式构成的。处于这种研究传统下的社会心理学强调了这些概念图式的社会决定性，因而强调了群体组织及其相关的社会氛围对个人观念的影响。

1930 年代期间，许多重要的格式塔理论家从纳粹德国逃亡到美国，其

中库尔特·卢因(Kurt Lewin)、雅各布·莫雷诺(Jacob Moreno,1925年移居到美国)和弗里茨·海德(Fritz Heider)是受格式塔影响的社会心理学的著名倡导者。卢因在麻省理工学院建立了一个研究中心,后来把它搬到了密西根,该中心成为研究社会观念和群体结构的重镇。另一方面,莫雷诺研究了应用心理疗法来揭示择友结构的可能性有多大。通过利用实验、控制观察和问卷法,他和同事们致力于探索如下问题:人们的群体关系在哪些方面向他们的行动,进而为他们的个人心理发展既提供限制性因素,又提供机会。尽管"社会计量的(sociometric)"这个词尤其与莫雷诺息息相关,但该词恰当地描述了来源于格式塔传统的常规的研究方式。

莫雷诺的工作深深地植根于针对人际关系的治疗导向,这反映了他早期在维也纳接受的医学和心理治疗方面的训练。他出版了一本重要著作(Moreno,1934),并在1937年创办一本杂志《社会计量学》(*Sociometry*)。在这些著述中,莫雷诺的目标是考察精神健康与他所说的"社会构型(social configuration)"是如何联系在一起的。这些构型是由人们参与其中的人际选择、吸引、厌恶、友谊以及其他关系等具体的模式带来的结果,这些模式也是大规模的"社会集合(aggregate)"如经济和国家等得以维持和再生产的基础。莫雷诺对小规模人际构型和大规模社会集合的区分清楚地表达了德国古典社会学的一些重要观念,特别是在韦伯、滕尼斯(Tonnies)①和齐美尔(Simmel)的著述中提出的观念。后者提出的"形式社会学"早就直接预设了许多社会计量学问题(Aron,1964;Simmel,1908)。

莫雷诺的主要创新之处是发明了"社群图(sociogram)",用它来表达社会构型的形式特征[1]。他认为,这些构型可用类似于空间几何中的图来表示,即用"点"代表个体,"线"代表个体之间的社会关系。现在看来,这种观点已经如此深入人心,以至于难以评价其在1930年代的新颖之处。在莫雷诺之前,人们已经谈到了关系"网"、"社会网络结构"等概念,有时甚至直接谈到关系的"网络",但是没有人致力于把这些隐喻式观念系统化为分析的图示。

对于莫雷诺来说,社会构型具有明确的、可区分的结构,利用社群图来图解这些结构,研究者可以把诸如信息从一个人传递到另外一个人以及人们之间相互影响的渠道等进行可视化处理。莫雷诺认为,社群图的构建允许研究者区分出领导者和孤立者,揭示出不对称性和互惠性,并用图来展示关联的渠道。他所使用的一个重要的社群图概念是社群"明星(star)",

---

① 原文应该为Toennies——译注。

即那位经常被他人选定,从而拥有极大的声望和领导地位的人。对于莫雷诺来说,"明星"这个概念指的是群体成员之间关系的易于可视化的图式。例如,在图 2.2 中,A 被群体内的所有人选择为朋友,而 A 仅仅选择了组内的 B 和 C 为朋友。所以 A 是组内具有吸引力的明星。

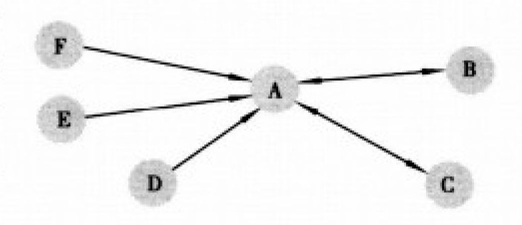


图 2.2　一个社群图:明星图

卢因(Lewin,1936)早期出版的一本书对群体行为进行了研究。该书的观点可概括为:群体行为是由行为所处的社会力量场决定的。他认为,一个社会群体存在于一个场中,场是由群体及其所处周围环境构成的社会"空间"。但是,群体的环境不应该被认为是完全外在于、独立于群体的因素。对群体成员起作用的环境称为臆想的环境(perceived environment)。它就是符号互动论传统中学者所说的"情境定义",其社会意义要根据群体成员对他们行动所处环境的认知和经验来进行积极的建构。因此,群体及其环境就构成了单个关系场中的要素。卢因认为,这个社会空间的结构特征可以利用拓扑学(topology)和集合论(set theory)中的数学技术来分析(Lewin,1959)。用数学术语来讲,"场论(field theory)"的目的就是探索在一个关系系统中群体及其环境的互依性,这个观点使得卢因接近于后来的一般系统论的发展(参见 Buckley,1967 有关这种研究框架在社会学中的应用的论述)。

拓扑学研究把社会场看成是由途径(paths)连着的点(points)构成的。如在社群图中一样,点代表个体、个体的目标或者行动等,而途径代表连接它们的互动或因果序列。因此,场论模型描述了社会构型中一些因果的、互动的互依性。在各个点之间运行的线把它们连接在一起,线的模式把一个场分为若干个离散的"域"。各个域之间由于缺乏沟通的线而相互分离:途径存在于每个域的内部,而不是各个域之间。个体在其社会世界中拥有的流动机会取决于其所处场内的各个不同域之间的边界。这些边界的限制性成为决定群体行为的一些"力量"。因此,全部社会场就是一类由各种力量构成的场,这些力量作用于群体的成员,并塑造着他们的行动和经验。

由海德(Heider)倡导的另外一支认知心理学,也为群体动力学理论的发展作出了重要贡献。他最初对态度和知觉进行社会心理学研究,并且特别关注一个人对待他人的各种态度如何达到一种"平衡"状态。个人拥有的各种态度只有在不产生心理张力状态的时候才会在他(她)的心里达到平衡。因此,心理平衡要依赖于个体拥有的各种观点互相不矛盾。海德尤

其关注人际关系的平衡,关注针对其他人的各种态度之间的和谐(或者不和谐)。例如,对于一个在情感上与其他两个人有密切关系的人来说,一旦这二人之间出现某种可察觉的冲突或敌意的时候,此人会作出怎样的反应,这是海德关注的问题。在这种情况下,整个态度场中出现了不平衡。海德(Heider,1946)认为,就最简单的情况而言,态度可分为积极态度和消极态度两类。当态度的符号都类似——都是"积极的"或者都是"消极的"时候,各个态度之间就存在"平衡"。例如,如果 A 喜欢 B,并且 B 喜欢 C,那么,只有当 A 也喜欢 C 的时候,才能达到平衡态。此时所有的态度都是积极的。需要着重注意的一点是,海德与卢因一样,认为这种分析与"从作为中心的个体角度来领悟世界"的方式有关:海德明显接受一种"现象学"的立场。从这种角度看,重要的问题不是 B 和 C 之间实际存在的关系,而是 A 对这种关系的(精确的或者别样的)臆想。"平衡"指的是一种心理学的、现象学的状态,而不是一个社会群体中任何实际存在的关系。

作为一种社会分析的理论框架,尽管场论被认为是一种智识的死胡同,但是卢因坚持的群体关系之数学模型却被认为提供了后续发展的坚实基础。在根据卢因的洞见基础上进行建构工作的学者当中,卡特赖特(Cartwright)居于尤其重要的地位。他和数学家哈拉里(Harary)共同开创了图论(graph theory)在群体行为方面的应用研究(Cartwright and Zander,1953;Harary and Norman,1953. 也参见 Bavelas,1950)。图论最初是由康内格(König,1936)提出来的,但是与 1930 年代在德国出版的许多著述一样,图论很少立即在智识世界中产生广泛的影响。直到 1950 年他的书在美国出版,并且哈拉里和诺曼(Harary and Norman,1953)推进了该书的思想的时候,图论思想相对于主流学术界的重要性才得以彰显。这些数学思想使得在群体动力学理论中出现重大突破成为可能。突破之处在于从个体心中的**认知平衡**转向社会群体中的**人际关系**(interpersonal)平衡。纽科姆(Newcomb,1953)是最早倡导这种转向的学者之一,他认为,两个关系密切的人有一种倾向,即他们对待第三个人或者事件的态度趋于一致。因此,研究者可以构建一个群体内部不同个体坚持的各种态度之间系统的互依性模型。后来,卡特赖特和哈拉里(Cartwright and Harary,1956)提出一种理论框架,推广了这一论断。在这些学者看来,卢因、莫雷诺和海德提出的洞见融在一起,达到一种新的更有说服力的综合(也参见 Harary et al., 1965,该书的准备工作自 1950 年代中期就开始了)。当然,利用数学来研究群体关系结构本身不是新观点,除卢因的研究工作之外,在 1940 年代末期还出现了利用多种数学模型进行研究而取得的成果(例如 Bavelas,1948;Festinger,1949)。然而,在卢因的研究基础上,卡特赖特、赞德(Zander)和哈拉里提出了关于群体凝聚力、社会压力、合作、权力

和领导方面的更有力的模型。

卡特赖特和哈拉里(Cartwright and Harary,1956)已经勾勒出了用线连点以便代表群体关系的基本思想——这也是莫雷诺的洞见。这样做的结果是产生了社群图或者"图",它代表了群体成员之间实际的人际关系网络,他们认为,可以用图论中的数学思想来分析这些关系。图论与许多人熟悉的数学中的变量图没有任何关系。一个图无非是由连接多个点的一系列线构成的,而图论是由一系列数学公理和公式组成的,这些公理和公式描述了由线连接的关系模式的性质。在卡特赖特和哈拉里的著述中,图中的点代表的是个体,线代表个体间的关系。一个图中的线可以被赋予一定的记号(+或者-),用来表示个体间的关系是"积极的"还是"消极的",也可以用箭头表示关系的"方向"。例如,赋予一条线以一定方向的目的是为了把A到B的方向和B到A的方向区分开:A到B的关系可能是积极的(A喜欢B),而B到A的关系可能是消极的(B讨厌A)。通过构造这种带"箭头"和"记号"的图,卡特赖特和哈拉里在分析各个个体的同时,也分析了群体的结构,而不仅仅从某个核心点的角度出发进行分析。因此,它向严格的社会学方向迈进了一大步。

通过考察无向图,就很容易理解卡特赖特和哈拉里提出的一些基本观点。在一个无向图中,A到B的关系被认为等同于B到A的关系。例如,这可能是因为人们的态度完全是相互的,或者因为他们共同参与同一个事件。因此,研究任何两点之间的线就不用考虑其方向。在一个无向图中,"均衡"描述的仅仅是构成该图的一些线上符号的特殊模式。例如,图2.3展示了三个行动者之间关系的三个不同图。在图(1)中,A和B相互之间有一个积极的关系,该图因而是平衡图,因为在A和C之间以及在B和C之间也存在着积极关系。然而在图(2)中,在A和C之间存在的消极关系给A和B之间的积极关系带来一种张力,因为在B和C之间也存在一种积极的关系。因此,该图是不平衡的。简而言之,如果我的朋友喜欢一个我讨厌的人,那么我和朋友之间的关系就会出现紧张。作为对这种状态的反应,我可能通过改变我自己对待他的态度,或者与朋友绝交等方式来劝说我的朋友,请他不要喜欢第三者。在一个非均衡的网络中,每个参与者都会处于同样的张力之中,因而都会试图削减体验到的张力[2]。群体关系处于动态流动之中,最终将出现均衡态——如果这是可达到的话——而这来源于牵涉到的所有参与者之间的行动和妥协。我们可用一个新图来表现那些目标在于恢复群体均衡的反应,该图的线上要带有不同的符号。例如,图(3)就代表了如下情况,即A成功地劝说B,使得B厌恶C,从而恢复了均衡态。

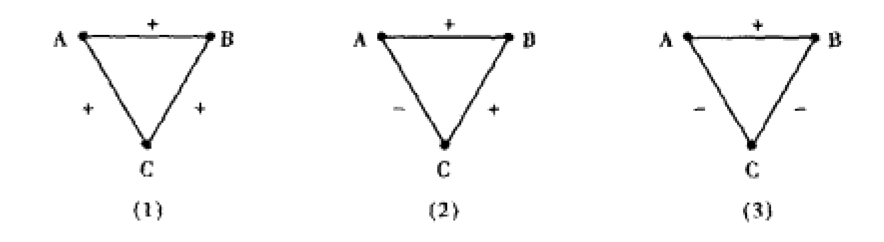


图 2.3　均衡结构和非均衡结构

卡特赖特和哈拉里认为,复杂的社会结构是由简单的结构构成的。特别是,它们是由多个如图 2.3 展示的那些互相重叠的"三人组(triad)"组成的。简单的三人结构是更大的社会结构的基础构件。他们认为,通过分析这些基础构件,就可以引申出复杂的社会关系网络的一些性质。例如,当构成一个网络的所有三人组都达到均衡的时候,该网络就是均衡的[3],这是最简单的情况。尽管均衡三人组思想可能是相当清楚并易于理解的,但是一个大的复杂网络也是"均衡"的,这种思想就稍难理解一些。确实,这种主张所提供的信息看起来既不令人感兴趣,又没有用处,然而,这种结论却是错误的。卡特赖特和哈拉里的研究工作的一个重大发现是,任何平衡图,无论它多么大或者多么复杂,都可以分为拥有令人感兴趣的特征的两个子图,每个子图**内部**的关系将是积极的,而子图**之间**的关系是消极的。这样,一个根据诸如团结关系来界定的平衡的社会网络就由相互之间存在着冲突和对抗的两个子图构成。

在一种简单的情况,即一个网络中的所有关系都是积极关系的情况下,其子群之一将是空集或者零集,即所有点都落入单个群体之中[4]。比较复杂的均衡网络并非如此,该网络分成的各个子群将突显出网络的一些重要结构特征。因此,把网络分解为各个子图仅仅是确定一个网络是否平衡的第一步。大多数涉及到网络平衡分析的数学工作都集中于试图发现这种分解的技巧。对一个均衡网络进行成功的分解,会使得研究者仅仅根据来自个体之间关系的信息就可以导出对网络结构的理解。这一发现对于理解群体结构来说有重要的意义。许多学者致力于探索,在什么情况下才可能走向针对网络的实质性分解技术,从而使得研究者区分出一个网络中存在的两个以上的子群。詹姆斯·戴维斯(James Davis,1967,1968)就是这方面的一位领军人物[5]。

在针对群体合作和领导的实验研究中,"平衡"概念特别有影响,它催生了一项关于自然情境下小群体行为的经典研究(Festinger,et al.,1959)。然而,许多来自于社会计量学传统中的小群体研究思想却被对一般系统论

以及控制论和理性行动理论感兴趣的学者吸收。确实，社会计量学思想最早在大规模社会系统中的应用恰恰是由这些学者开创的。最初的研究探讨了疾病通过接触链在人与人之间的传播，其目的是推导出可预测的关于传染的流行病学模型。这种研究的领袖人物是拉波波特(Rapoport)，他探讨了此类研究的含义(Rapoport,1952,1958)，并激发了人们的兴趣，使人们利用这种思想来研究思想观念和创新事物的传播。尽管此前已经有过这类研究，即人们已经研究过谎言、流言蜚语的传播等，但是直到20世纪60年代，利用一些网络概念于此类研究的重要成果才得以出现(Coleman et al.,1966;Fararo and Sunshine,1964)。

人际结构和派系

前文已经指出，在社会计量学传统下的理论工作倾全力于发现一些将网络分解为子群体的方法。对被称为"派系"、"聚类(clusters)"或者"块(block)"等各种术语的研究也是20世纪30、40年代哈佛大学研究传统的一个特点。在这种研究中，对大系统中"非正式关系"的研究导致了一个经验发现，即这些系统实际上包含很多凝聚子群体。因此，如果可以得到某一社会系统中的关系数据，那么研究者面临的任务便是去发现一些技术，用它们来展示该社会系统中的子群结构。这一任务现在仅仅部分解决了。

拉德克利夫-布朗和涂尔干(Durkheim)对这一研究传统的影响很大。拉德克利夫-布朗的观点在澳大利亚的人类学家中颇具影响，因为他曾经在该国执教多年。在W·劳埃德·沃纳(W. Lloyd Warner)的研究工作中，拉德克利夫-布朗的影响特别大，前者于1929年进入哈佛大学，与他的澳大利亚同事、心理学家埃尔顿·梅奥(Elton Mayo)一起工作。此二人合作开展了一系列针对美国的工厂、社区生活的研究，这些研究息息相关，并被他们看成是拉德克利夫-布朗所关注的结构在这些领域中的应用。

1926年，梅奥从澳大利亚移民到哈佛，目标是在哈佛新组建的商学院的一项研究计划中担当主角。主要受到生物学家L·J·亨德森(Henderson)的重要影响，梅奥才接触到社会学思想，前者在哈佛的同事中积极推广帕累托(Pareto)的思想。亨德森认为，帕累托的研究是一门真正科学的社会学的唯一恰当的基础，进一步说也是用来反对革命性的马克思主义的唯一可行的政治堡垒。梅奥逐渐意识到帕累托所谓的行动的一些"非理性"因素，他对个人动机的心理学关注也因而得到补充。对梅奥来

说，经济行动并不完全是一种理性的行动形式，它也被诸如群体团结等很多非理性的情绪因素所结构化。帕累托也是一位精英论的伟大倡导者，梅奥认识到，如果一个管理精英认识到群体关系对经济动机的影响，那么他就会最成功地控制工人的行为。作为一位恰当的训练有素的现场研究者，沃纳对哈佛研究纲领的贡献在于，他对见于某些特定社会场景中群体行为的实际模式给予极大的关注。沃纳向梅奥所关注的理论和应用方面输入了一种经验维度。尽管有这些差异，或者说正因为如此，二人在哈佛开创的研究工作才对于社会网络分析的发展来说是尤其重要的。虽然他们在那里共事仅六年时间，但是事实表明他们的研究极具影响力。他们与同事们共同开展的一些主要计划是对芝加哥的霍桑电气工厂（Hawthorne electrical factory）和新英格兰[①]（New England）的一个被称为“扬基城（Yankee City）”的社区进行研究。

霍桑研究已经成为社会研究的经典，这里只需稍加论述（参见 Rose，1975 的有益的讨论）。简单地讲，这是于 1920 年代由芝加哥西部电气公司霍桑工厂的管理者开创的有关工人效率的一系列研究。这些管理者试图揭示，工作方面的自然条件（如温度、光照、休息时间的长短等）的改变是如何影响生产率的。他们惊奇地发现，生产力的增长与特定物质条件的改变几乎无关。为了试图理解这一悖论式的结果，管理者们拜访了梅奥及其哈佛小组，请这些学者指导他们重新设计研究计划。梅奥的结论是，带来生产力增长的关键因素恰恰是工人们参与了研究计划这个事实：工人们很兴奋，因为管理者关注他们，并且正是由于他们意识到自己参与并整合到工厂生活中，才促进他们努力工作。

在沃纳的建议下，霍桑的研究者们开始了一项“人类学”研究，在工厂的自然环境下对劳动小组的行为进行观察。他们的观察场地是触排布线室（bank wring room），其研究方法与人类学家在一个异化社会的乡村中进行的现场研究类似。他们尽可能记载了所有可能观察到的群体行为，目标是建构一个完全人类学意义上的解释。在社会网络分析的发展过程中，霍桑研究的特殊重要性在于，研究者用社群图来报告群体结构。就像在一个乡村社区中的亲属关系结构可以用一个谱系图来展示一样，霍桑小组也构建了很多社群图，用来展示工作小组内部的非正式关系。

在霍桑实验的主报告（Roethlisberger and Dickson，1939：555 及以后）[6]

---

① 新英格兰（New England）是美国东北部地区缅因（Maine）、佛蒙特（Vermont）、新罕布什尔（New Hampshire）、马萨诸塞（Massachusetts）、罗德艾兰（Rhode Island）和康涅狄格（Connecticut）这六个州的总称——译注。

中包含了研究小组构建的各种社群图。他们认为，这些图反映了触排布线室中工人的“非正式组织”，而不是在管理组织图中描述的正式组织。他们构建的社群图表达了群体行为的每个方面，诸如参与游戏、隔着打开的窗户争吵、工作对换、帮助、友谊和敌视等。霍桑研究是运用社群图来描述现实情境中观察到的实际关系的首次研究。在他们画的图中，用圆圈代替人，用箭头表示关系。这些图与后来群体动力学研究者构造的社群图明显相似，但是，研究者并没有指明他们是如何想到利用此类图的。例如，他们对正在形成中的莫雷诺的思想避而不谈。然而，从图 2.4 中可以看到，图不仅可以代表管理者所使用的正式组织图，也可以代表工厂的一种非常熟悉的特征，即电气布线图。我们必须假定，沃纳的影响激励了这些研究者，使他们通过接受组织情境中的这些影响，进而适应传统的人类学亲属关系图。

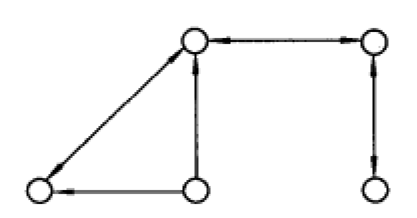


图 2.4　一个霍桑社群图

在画触排布线室的社群图的时候，需要遵循一般的约定，但是这些约定都是艺术性的而不是社会学的。每页上的各个圈的具体位置都由画手自己决定，主要的限制性条件仅仅在于，观察者确定下来的任何一个子群中的各个成员应该画得尽可能地接近。除此之外，清晰性、简单性这些纯艺术原则也主导着图画的设计：例如，交叉的线数应该尽可能少，线的长度不应该相差太大。研究者确定下来的子群——他们称之为“派系(cliques)”——指的是如下这些工人群体，即工人们认为自己是他们所处环境的重要成分。很多人类学家一般利用“当地人”的范畴和概念作为进入群体生活的结构性质的指示器，与此很相像，工人们自己的术语也被用作“派系”存在的指标。在对群体的行为和群体所使用词汇的观察中，区分出“前台群体”和“后台群体”，并把二者看成是触排布线群体(bank wring group)内部的子群体。他们并没有试图利用社群图本身来区分在社群图意义上界定的“派系”；只是把根据在社会上理解的子群画在社群图中[7]。然而，通过这种方式画出群体结构之后，研究者就很少进一步使用社群图了。社会网络如何型塑个体的行为，他们对此似乎缺乏任何理论方面的理解。

与此同时，沃纳开始对美国东北部地区的一个名为纽伯里波特①

① 美国马萨诸塞州北部的一座城市——译注。

(Newbury port)的小城进行研究,他起了一个学名:"扬基城(Yankee city)"。这一研究是在1930年到1935年间进行的,该研究也被认为是关于现代城市社区的一项成熟的人类学研究。该研究综合运用了观察、访谈以及历史文献法。然而,在此研究主要阶段结束之际恰逢沃纳和梅奥之间渐增抵牾之时,沃纳因而跳离哈佛,就职于芝加哥大学,此时他的导师拉德克利夫-布朗已经是芝加哥大学的一位访问教授了。沃纳和拉德克利夫-布朗在芝加哥共事2年时间,这是一个对来自于扬基城的研究资料进行最紧张分析的阶段。此后,沃纳一直在芝加哥供职,在此指导并发起了许多相关的研究,其中最著名的是对美国东南部地区(Deep South)[8]的"旧城(Old City)"的研究。

沃纳的早期工作已经运用传统意义上的涂尔干和拉德克利夫-布朗方法和思想来研究一个澳大利亚部落,正是通过与梅奥的接触才使沃纳首次形成了运用人类学方法来研究一个现代城市社区的思想。沃纳最初希望研究芝加哥的街区,即霍桑实验地所在的街区,但是芝加哥社会学派的研究迫使他得出结论认为这些街区是"无组织的",因而经不起人类学的检验(Park et al.,1925)。沃纳感觉到,只有在新英格兰(New England)和南部的一些州,他才能发现自己希望加以研究的发展完善并且整合的社区。

沃纳的工作具有多方面的理论影响。尽管拉德克利夫-布朗对沃纳的影响最主要,但是沃纳还是把这种影响与关于社会的一种有机的、系统的模型结合起来。毫无疑问,这种模型又受到了亨德森对帕累托的解释的影响。这促使沃纳强调一个社区在结构化过程中的某些因素,如稳定性、凝聚力和整合性等。但是,他也吸收了齐美尔关于"互惠关系"和"数字对群体生活有影响"的思想①。我已经指出,正是齐美尔(Simmel,1908)才开创了作为社会生活之基石的二人群体(dyads)和三人群体(triads)的研究。沃纳延续了齐美尔和其他德国社会学家的术语,同时也接受莫雷诺的用语,进而谈到社会构型,认为一个社区的社会组织是由人与人之间据以互动的关系网构成的。

沃纳指出,构成一个现代社区的社会构型是由许多子群,如家庭、教堂、班级和协会等组成的。除此之外他们还发现另外一类子群,他称之为"派系":它是一个非正式群体,其中的成员都有一定的群体感和亲密性,并在群体内部建立起某些群体行为规范(Warner and Lunt,1941:32)。派系是"一个关系紧密的非亲属群体,其成员数多少不一,从2人到30人或者

① 齐美尔认为,群体成员的数目对群体的生活有重要影响——译注。

更多"（Warner and Lunt,1941:110）[9]。因此，对沃纳来说，在社区研究中，派系的重要社会作用就相当于非正式群体在霍桑实验中的重要作用一样。这个概念描述了非正式人际关系中的一种特定的构型。

扬基城的研究者声称，在该城市中能找到数目极多的此类派系。扬基城的许多回答者用诸如"我们一伙的"、"我们圈内的"等词汇来指代他们所说的群体，派系主要就是这些群体。根据被研究者作出的评论，研究者发现了这些派系的存在。此后，沃纳及其同事认为，相对于把个人置身于社会中的家庭来说，派系仅居于次要地位。人们是通过家庭和派系成员这种"非正式的"和"个人的"关系，而不是简单地通过经济系统和政治系统中的"正式"关系整合到社区中的。任何人都可以成为许多派系的一个成员，并且"这种派系成员的重叠会进一步扩展为一个相互关联的网络，该网络把一个社区内部的几乎全部人口整合到一个巨大的派系关系系统之中"（Warner and Lunt,1941:111）。毫无疑问，利用网络术语来描述整个社会被建构成为各个子群体的过程，这种描述如果不是最早的一个应用，至少也是最早的应用之一。

扬基城报告运用各种图表来对诸如阶级结构和家庭组织等进行模型化表征，毋庸置疑，他们同时也构造了派系图。为了表征所描述的社会结构，他们把派系绘图成为在一个维恩图（Venn diagram）中相互交织的一系列圆圈（Warner and Lunt,1941:113），但是他们没有进一步对这些图进行任何正式的结构分析。然而，在扬基城报告的第二卷中，就有进行现在所谓的"位置分析（positional analysis）"的尝试了（Warner and Lunt,1942:52，图 X）。他们提出了一系列矩阵，用以展示占据每个结构位置的人数。图2.5 就表示其中的一种形式。在区分出扬基城中的6 类等级和31 类派系之后，沃纳和伦特（Lunt）用一个数据矩阵对等级和派系成员进行了交叉分类。每一类派系要根据其全体成员的主要阶层成分来定义，矩阵的各个元素中的位置内容表明每一阶层中属于各个派系的人数[10]。在这些大量可能的组合，即6 乘以31 等于186 个组合中，他们认为实际发生的"位置"只有73 个，矩阵中其余的113 个位置都是空的。通过构建类似的关于等级和其他社会群体（如各种正式关系，各种家庭等）之间关系的矩阵，他们能够把各种矩阵结合、叠加在一起，从而在总体的组合网络中区分出89

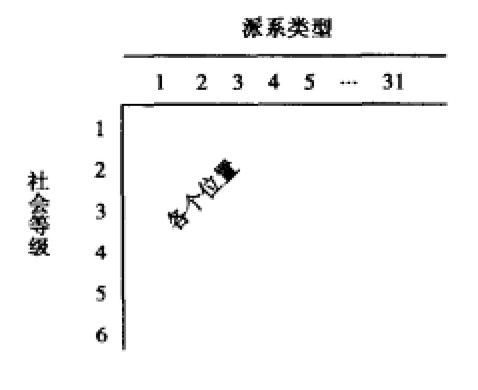


图2.5　派系矩阵

个位置[11]。他们使用的特定方法比较繁琐,我们也无需深入到这种早已过时了的操作细节中,但是扬基城的研究工作仍然具有吸引力,因为它试图开创这种正式的结构分析方法。

沃纳的同事于1936年在美国东南部的"旧城(Old City)"开始另一项研究,从而深化了"派系"(Davis et al.,1941)思想。在考察"旧城"中"多彩的社会"时,他们按照沃纳的方法,把派系看成是相互交叉的各个圆圈,并在一个由等级和年龄界定的空间中,勾画出那些最积极派系中叠加的成员(Davis et al.,1941:213,图12)。他们称之为"社会空间"及其"两个维度",但是他们没有明确提到卢因关于拓扑式的田野模型的任何研究工作。这种研究的主要创新之处在于它试图探究派系的内在结构。研究者认为,可以把一个派系看成是由如下三个"层次"组成的:"核心(core)",由最经常、最紧密地聚合在一起的人组成;"初级圈(primary circle)",由那些时常共同参与的核心成员,但本身从来也不能构成一个群体的人组成;次级圈(secondary circle),由那些仅仅偶尔参与其中,因而"几乎不是其成员"的人组成。在对60个派系进行分析的基础上,利用类似于扬基城研究中所使用的技术,他们提出了关于派系**之间**联系的大量结构假设。例如,他们认为一个派系的外围的、低层次的成员只有通过本派系的高级核心成员才能够接触到其他派系的高级成员。

在霍桑、扬基城和旧城研究中出现的思想在发展过程中与那些在小群体研究中的社群图传统遥相呼应,但是在1930年代和1940年代,没有证据表明这两种研究传统的代表人物相互意识到对方的研究工作的存在。然而,在乔治·霍曼斯(George Homans)的著述中出现了这两种研究传统的第一次重要的交叉。作为哈佛大学社会学系的一位职员,霍曼斯不满意于一些同事如帕森斯(Parsons)的巨型理论(grand theory),认为这种理论是在非常抽象的分析层次上运作的。霍曼斯感觉到,社会理论必须建立在对小规模社会互动的确切理解的基础上。为了达到此目的,在1940年代末,他开始尝试对当时在美国已经进行的大量有关小群体的研究进行综合。他的目的无非是通过吸收社会心理学家的经验研究和社会学家、人类学家的观察研究,从而达到理论上的综合。这种综合围绕着如下观点进行,即各种人类行动使人与人之间出现了互动,这些互动在"频次"、"持续的时间"和"方向"[12]上各不相同,并且互动是人与人之间产生"感情"的基础。霍曼斯把莫雷诺的社会计量学看成是为利用这种理论于一些特定的社会场景提供一种方法论框架。为了展示自己的观点,霍曼斯重新考察了大量早期的研究工作。

在《旧城报告》中,有一部分至少在社会网络分析领域获得了极大的名

誉，因为该部分由霍曼斯作了重新的分析。在该部分中，戴维斯和他的同事应用矩阵方法来分析 18 位女性参与 14 个社会事件的情况[13]（Davis et al.，1941：第七章）。霍曼斯得到这些数据，用矩阵加以表征，并第一次正式发表社会网络分析“重排矩阵”方法（也可参见 Festinger，1949）。“旧城”矩阵有 18 行（女性）14 列（事件），矩阵中的“x”项代表某位女性参与了一个特定的事件。霍曼斯认为，初始阵无需根据什么重要的顺序加以排列——例如，各个列简单地根据事件发生的日期次序进行排列即可。因此，所分配的各个交叉值在矩阵中是随机排列的。他相信，对矩阵的各行和各列进行重排将把特定的某些女性重点参与的事件集中在一起，这种重排就可以揭示派系的一些重要结构特征。他对此方法描述如下：

> “我们把代表多数女性参与的事件的各列置于核心地位……，把代表少数女性参与的事件的各列放在边缘。就所涉及的这些行（列）来说，我们会把那些代表最经常一起参与社会事件的妇女的行放在顶部或底部。在任何模式出现之前，需要做大量此类的重排工作。”（Homans，1951：83）

霍曼斯认为，这种重排必须持续下去，直到各交叉值的分布能够展示出清晰的模式，他也制作了一个重排阵，其中清楚地标示着女性出现的两个“派系”：在重排阵的交叉项中出现了两个明显不同的群体。霍曼斯的方法与后来所谓的“块模型（block modelling）”方法类似，但是他没有利用任何正式的数学方法。实际上，他的重排似乎仅仅是一种试错法（trial-and-error process），该方法要持续下去直到他能够清楚地看到一种明显模式的出现。

图 2.6 简化地表达了霍曼斯所执行的再分析。这两个矩阵表达了一个假想的 8 人参与 8 个事件的情况。在矩阵（1）中，各个“x”项平均地散布

（1）初始矩阵

		事件							
		1	2	3	4	5	6	7	8
人	Ann	×		×		×		×	
	Beth		×		×		×		×
	Chris	×		×		×		×	
	Don		×		×		×		×
	Ed	×		×		×		×	
	Flo		×		×		×		×
	Gill	×		×		×		×	
	Hal		×		×		×		×

（2）重排矩阵

		事件							
		1	3	5	7	2	4	6	8
人	Ann	×	×	×	×				
	Chris	×	×	×	×				
	Ed	×	×	×	×				
	Gill	×	×	×	×				
	Beth					×	×	×	×
	Don					×	×	×	×
	Flo					×	×	×	×
	Hal					×	×	×	×

图 2.6 矩阵的重排

在整个矩阵之中，但是对该矩阵的行和列按照矩阵(2)展示的顺序进行重新排列之后，就带来两个不同的子群之间的结构对立：Ann，Chris，Ed 和 Gill 一起参与了事件 1，3，5，7，而 Don，Flo，Beth 和 Hal 则共同参与了事件 2，4，6，8。这里存在两个集合的人和两类事件。我们可以意识到，通过这种试错法进行的重排不会是如此容易的事情，即使对于像该矩阵这样的相对较小，并且数据不很紧密地组织在一起的矩阵来说也不容易。实际的数据是关于 18 位女性参与 14 个事件的数据，它要花费大量时间来分析。不仅如此，我们也无法保证霍曼斯重排的最终结果与任何其他人的重排结果是一样的，因为不存在一个“正确的”、用来区分结果的标准。正是由于这些原因，后来很多此类分析都涉及找出一种程序算法，从而可利用计算机来执行这种重排的工作了。

为了进一步说明他的立场，霍曼斯重新分析了霍桑实验的触排布线室数据。利用观察者构建的一些社群图，他考察了洛特里斯伯格(Roethlisberger)和迪克森(Dickson)区分出来的派系(Homans，1951：66-70)。但是霍曼斯保留了这些初始的派系之分，并没有秉承他自己分析旧城数据的路线，即没有试图对派系的结构进行社会计量学的分析。然而，他确实暗含地指出(尽管没有明确指明)，矩阵重排法已经被最初的霍桑研究者所使用(Homans，1951：84)[14]。

霍曼斯构建的用来解释群体行为的理论框架是对早期小群体研究者的模型的进一步阐释，该模型把群体理解为处于某一环境中的系统。他把任何群体的结构都分为“内在系统”和“外在系统”两种。前者表达了通过群体成员之间的互动所产生的情感，而群体行动要通过后者才与环境的适应问题建立联系[15]。环境本身是由群体行为的物理的、技术的和社会的环境组成的。霍曼斯主要关注的是内在系统，并且把它看成是一个比该概念所归属的“非正式组织”概念更科学的概念。因此，他的兴趣是通过把来自于非正式组织研究的洞见转换为关于内在系统之结构的命题，从而对此进行科学的探究。

为了达到这个目的，他就内在系统提出了大量假设。首先假定：相互经常接触的人之间也倾向于相互喜欢，互动次数越多，相互喜欢的程度越大。在外部系统中，由于存在着诸如指导者和管理者强加的一些要求这样的环境限制性因素，因此，外在系统中存在的多次互动将使得工作组中的成员趋于发展出喜欢性的情感，并将共同参与更深层次的互动，而这与外在系统的需要是无关的。他论证道，内在系统正是通过这种方法才进入到可分解为多个派系的复杂社会构型之中的[16]。

尽管霍曼斯对社会学和人类学研究的理论综合具有解释力，但是他的

工作几乎没有直接激发出什么重要进展。霍曼斯本人也越来越关注用行为主义和理性选择模型来解释社会行为,从而与“交换论”渐趋一致了(Homans,1961)。霍曼斯的一位同事罗伯特·贝尔斯(Robert Bales)开展了某种有趣的小群体研究(Bales,1950),但是他没有利用社会计量学方法于自己的研究之中,而是越来越与帕森斯的结构功能主义(Parson et al.,1953)关联在一起。许多对均衡观点有贡献的学者的研究工作完全返回到心理学的关切上来,并且在引导这些研究者返回到关于情感的社会心理学方面,费斯廷格(Festinger,1957)的有影响的著作成为一种重要的纲领性声明。群体动力学领域几乎停滞下来,大多数进展都是关于均衡、派系和聚类的纯数学问题,尽管这些数学解释成为后来哈里森·怀特的推进性研究的重要且丰富的来源,它们对于20世纪五六十年代社会研究的形成的影响仍然甚小。

网络:全网与局部网

促使社会网络分析的框架出现新转折的是曼彻斯特大学社会人类学系的一个研究小组,尤其是约翰·巴恩斯(John Barnes)、克莱德·米切尔(Clyde Mitchell)和伊丽莎白·博特(Elizabeth Bott)[17]等人的研究工作。曼彻斯特的人类学家们甚至比哈佛的同行们更受到拉德克利夫-布朗的重要影响,他们力图在一个新方向上提出自己的观点。他们强调变迁和矛盾,而不是整合和凝聚。其核心人物是马克斯·格卢克曼(Max Gluckman),他把对复杂的非洲社会的兴趣与发展一种结构视角结合起来,这种结构视角认识到,冲突和权力在社会结构的维持和转型中都扮演重要角色。对于格卢克曼来说,冲突和权力是任何社会结构的内在成分,他的分析强调了在社会整合过程中始终存在的协商、讨价还价和强制等活动。格卢克曼积极地鼓励他的那些正在进行小规模人际社区研究的同事和学生来从事这些主题研究。

在1950年代,帕森斯式的社会学研究以及人类学中的文化研究占据着主导地位,这对于把曼彻斯特学派引向一种独具特色的批判传统来说是一个重要因素。经典社会学家强调把行动放在其所在社会关系结构中的位置来理解,而帕森斯认为,必须把行动解释为对内化的价值取向的表达。曼彻斯特人类学家强调把结构看成是关系网络,所以他们的研究工作就把社会网络分析的形式技术与一些重要的社会学概念结合起来。结果表明,

这是一种给人印象深刻的、强有力的结合，它与当时社会学中正在出现的冲突论框架很接近，但是他们对人际关系的强调意味着它并不是以一种完全代替帕森斯式理论的面貌出现的。正因为如此，社会网络分析只能被看成是一种特殊的研究方法，而不是对传统社会学的批判式替代。

随后，曼彻斯特的学者较少关注一个社会的正式制度化规范和体制，却更加关注由冲突和权力的运作引起的实际存在的关系结构。来自于过去并且目的在于理解简单的、以亲属为基础的社会的一些理论观点不能解释这些现象。正是由于意识到这种不足，他们才开始对诸如拉德克利夫-布朗等学者所提出的社会关系"网(web)"、"网络(network)"等这些隐喻式概念进行系统化分析。

最初，这些研究者只是在隐喻的意义上使用社会网络观念。但是，在1950年代早期，巴恩斯率先在一种比较严格的、分析的意义上使用网络观念。他的研究对博特产生重要影响，随后二人紧紧围绕在社会计量学传统中曾经进行过的研究工作开展合作研究。他们发表的各类论文(Barnes，1954；Bott，1955，1956)受到社会人类学家的广泛欢迎，而社会网络这个概念也似乎满足一个用来理解复杂社会的恰当概念之需要。西格弗里德·纳德尔(Siegfried Nadel)在一系列演讲和相关的一本书中采纳了这种研究(Nadel，1957)，这本书成为来自该领域的一位重要人物的纲领性宪章。然而，完成了纳德尔提出的任务并提出社会网络分析的系统框架的学者是克莱德·米切尔。米切尔转向来自于早期社会计量学研究的图论数学，并把这些数学技术看成是一种独特的**社会学**框架的基础。他和同事(Mitchell，1969)的著述总结了成型于1950年代的这些思想，提出了一整套社会学概念，并相信这些概念能够充分地把握社会组织的**结构**特征。米切尔把图论和社会计量学翻译成一种社会学框架，这一做法激发他关注曾经由梅奥、沃纳和霍曼斯提出的"非正式"组织和人际组织的那些特征。

巴恩斯在中非的罗得斯-利文斯通[①]研究所(Rhodes-Livingstone Institute)开始自己的学术生涯，这个研究所是许多曼彻斯特人类学家的一个主要研究中心。巴恩斯于1949年加入曼彻斯特大学，随后决定对挪威西南部的一个渔村布雷姆斯(Bremnes)进行现场考察。这是一个孤立的小村，其成员几乎完全通过其成员之间的亲属关系连接在一起。该村是一个复杂、分化的社会的有机成分，有其内部的经济、政治和其他制度，这些制

① 该研究所可能是以如下两个人命名的。罗得斯(Cecil Rhodes，1853—1902)，生于英国的南非殖民地政治家；利文斯通(David Livingstone，1746—1813)是一位英国苏格兰的传教士、非洲探险家，是第一位看到赞比亚河及维多利亚瀑布的欧洲人——译注。

度没有被很好地协调到一个整合系统之中。巴恩斯尤其关注在社区整合的生产过程中亲属、朋友和邻里所起的作用。这些原初的关系并不直接与地域场所或者正式的经济、政治结构关联在一起。恰恰相反，它们形成了一个独特的、相对整合的非正式人际关系域。巴恩斯声称，"全部社会生活"可被看成是"一个集合的点，其中一些点由线连在一起"形成一个关系"全网"。在这个全网中，人际关系的非正式域可以看成是其中的一部分，即一个"局部网(partial network)"。

博特是一位加拿大心理学家，她在芝加哥大学的劳埃德·沃纳指导下研究人类学，并且可以想象，她与巴恩斯一样对扬基城的研究比较熟悉。1950年，她加入塔维斯托克研究所(Tavistock Institute)，不久就对英国的一些家庭进行现场研究。博特主要关注这些家庭的亲属关系，她用"网络"这个概念作为一个分析工具，用它来探究在亲属关系中表现出来的各种不同模式。其研究成果发表在两篇有影响的论文和一本书(Bott，1955，1956，1957)中，这也是博特于1956年获得伦敦经济学院(London School of Economics，LSE)博士学位的基础。

毫无疑问，在她的研究中逐渐形成的理论框架受到她本人在塔维斯托克研究所的一些同事的影响，这些同事在1947年就加入了位于安阿伯[①](Ann Arbor)的群体动力学研究中心，并出版了《人类关系》(*Human Relations*)这本杂志。作为一位对心理疗法(Psychotherapy)感兴趣的心理学家，博特注意到了莫雷诺的研究工作。确实，博特和巴恩斯都在自己的论文中引用过莫雷诺的成果。然而，对博特的研究产生过直接影响的是卢因的场论，甚至巴恩斯也探讨过在布雷姆斯社会中多种独特的活动场的存在。卢因、费斯廷格、纽科姆、卡特赖特和其他美国小群体研究的领军人物在《人类关系》杂志上发表很多论文，也正是在这本杂志上，博特和巴恩斯开始发表关于社会网络的研究成果。

巴恩斯最初于1953年在曼彻斯特和牛津的讨论班上提出自己的思想。1954年，博特了解到巴恩斯的工作，接受了"网络"这个术语并把它作为自己的理论解释的基础。巴恩斯发表论文之时正在伦敦经济学院的雷蒙德·弗斯[②](Raymond Firth)手下工作。这一年博特也已经注册攻读哲学学位，她在伦敦经济学院和曼彻斯特两处都汇报过自己的研究工作的初

---

① Ann Arbor(安阿伯)，美国密西根州东南部城市，密西根大学所在地——译注。

② 雷蒙德·弗斯爵士(Sir Raymond Firth，1901—2002)，英国社会人类学之父，功能学派的主要代表人物之一，是马林诺夫斯基的学生和学术继承人，也是费孝通教授的老师，以研究大洋洲的民族学著作而闻名，曾任伦敦经济学院和剑桥大学教授——译注。

稿。完全出于是文物工作者的原因,这里没有给出这些传记性的细节,也不仅仅是为了展示学术网络之重要性。我的关注点在于,展示少量的关键人士是如何在很短时间内肩负起建构一种英国社会人类学理论创新之基础的责任的。一旦巴恩斯和博特取得突破,就打开了进一步发展的大门,并在来自美国学者的进一步训练的基础上进一步强化这些进展。

在赋予这种导向的理论进展以合法化地位的过程中,西格弗里德·纳德尔是一位重要人物。他是一位关注社会网络分析的奥地利心理学家。受科勒和卢因的影响,他于1930年代转向人类学研究。1955年,纳德尔在伦敦经济学院作了一系列有关社会结构的讲演。巴恩斯和博特的工作无疑对他的研究工作的进展以重大影响,而在这些讲演的正式出版物(Nadel,1957)的前言中,纳德尔也提到把二人看成是自己的评议人和朋友。纳德尔首先把"结构"定义为构成一个整体的多个要素的结合或排列。他区分了结构的**形式**和**内容**,认为这样就可以利用比较法描述结构的一般特征。为了达到建构正式模型的目的,他倡导对结构进行数学研究。

根据纳德尔(Nadel,1957:2)的见解;社会结构是关于关系的"一个整体的系统、网络或者模式",是分析者从具体的、可观察的各种个体行动中抽象出来的。他认为,"网络"意指"各种关系的交织,其中一种关系中的互动决定着其他关系中的互动"(Nadel,1957:16)。纳德尔则特别指出,"角色"应该被看成是社会学理论的核心概念。社会结构是"角色"的结构,而角色及角色丛是通过多种互依活动之网来界定的。他认为,在角色分析中应该用代数、矩阵方法,但除了一两处简单的提示之外,他没指明具体如何运用这些方法。由于英年早逝(1956年去世),他没能进一步推进自己提出的设想。

在20世纪五六十年代,米切尔及其他与曼彻斯特和罗得斯-利文斯通研究所有关的学者试图系统化这种观念。事实上,米切尔是纳德尔的抱负的真正传人。1969年,米切尔把社会网络进行符号化处理,这就扩展了巴恩斯的关于人际关系域的概念,把它变成"人类的秩序(personal order)"[18]。"人类的秩序"指的是"多个人与一群人之间的个人关联以及这群人自身之间的关系"模式(Mitchell,1969:10)。对米切尔来说,这种互动模式是网络分析的领域。他进一步指出,这种人际网建立在行动的如下两个不同的理想型基础之上。首先是"交往",它涉及到信息在个体之间的传递、社会规范的建立以及一种程度共识的达成等。另一方面,存在一种"工具性"或者目的性的行动,它涉及到人们之间的物质和服务的交换(Mitchell,1969:36~39)[19]。这两种理想型以多样的方式结合在一起,从而形成了一些具体的互动之网。任何一种特定的行动都将把这两种理想型中的一些要素结

合起来,因此,特定的社会网络将既体现出信息的流动,也体现出资源和服务的传递。

米切尔继续把一个社会的“全网”概念化为“在任何社区或者组织的边界之内或超出其外的始终交织、始终分叉的关联丛”(Mitchell,1969:12)。他认为,在实际的研究中,我们总是必须选择整体网的某些特定的方面予以关注,他把这些方面概念化为“局部网(partial network)”。这种抽象过程建立在如下两点基础之上,尽管米切尔自己几乎完全关注其中之一。首先,一种抽象概念关注的是围绕着某一特定个体的关系,从而产生各种社会关系的“‘个体中心’网(‘ego-centered’ networks)”。其次,另外一种抽象观念关注的是相对于社会活动的某一特定方面网络的总体的、“全面”的特征,如政治关联、亲属义务、朋友或工作关系等。对于米切尔和大多数曼彻斯特的研究者来说,围绕着个体的局部网是他们关注的焦点。在此类研究中需要区分出个体,追溯他们与他者之间的直接和间接的关系。这种研究将产生一系列个体中心网,这些网都是关于所研究的每个个体的网络。与之类似的一种研究是由博特在早期分别针对夫妻的个体中心网的研究中给出的,她测量了这些网的“关连性”,也测量了婚姻伴侣的网络之间的重叠程度。

米切尔认识到上述第二类抽象观念——根据“内容”和关系的意义来界定局部网——的重要性,但是他认为这种网络也要围绕着一些特定的行动者。社会学家和社会人类学家所研究的“局部网”总是那些关注特定类型社会关系的个体中心网。米切尔指出,多数此类网络都是“多线条”或“多丛”的,它们涉及到大量有意义的各类关系的结合。这样看来,巴恩斯最初提出来的网络概念以及被博特接受的网络概念,就是一种局部网络,在这种网络中,亲属、朋友和邻里等被组合成为一种单一的、多线条的关系,从而不能轻易拆分为其各种构成成分。

米切尔论证道,人际关系网络可以通过大量的概念来分析,这些概念描述了所涉及到的关系的性质。这些概念包括关系的“互惠性(reciprocity)”、“紧密性(intensity)”、“持续性(durability)”等(Mitchell,1969:24-9),它们反映了霍曼斯对方向、频次和紧密性的区分。某些(但不是全部)关系涉及到交易或交换,因而可以看成是从一个人“指向”另一个人。因此,这种关系的一个重要测度是交易(或者定向)在多大程度上互惠。例如,一个人可能选择另外一个人作为自己的朋友,但是这种选择可能不会得到回选:该选择者可能被忽视或被一脚蹬开。多线关系能够包含一种由互惠关系和非互惠关系相互补偿构成的均衡。例如,通过这些关系,资金支持可能沿着一个方向流动,而政治支持则取另外一个方向[20]。

"持续性"测量的是在特定交易中激发的内在的关系和义务是如何持续下去的(Mitchell 引自 Katz,1966)。在互动中稳定地活动的关系和义务等具有高度可持续性,而那些只在一两次活动中保持的关系等是高度可变的。例如,亲属义务具有很高的持续性,它们一般持续人的一生,而那些出于一种特定的限定性目的的关系则容易转瞬即逝。"紧密性"指的是在一种关系中涉及到的义务的强度。它或者反映了投入到这些义务中的程度,或者反映关系的多元度:多线条的关系有紧密的倾向,因为它们有传播性[21]。

米切尔追加了一些概念,这些概念来源于把图论翻译为社会学语言,从而可用来描述社会网络之结构。例如,他把"密度"这个概念看成是网络的完备性(completeness):所有可能出现的关系实际上在多大程度上出现。这个概念就是巴恩斯和博特曾经试图用他们自己提出来的"网状物(mesh)"和"关联度"所描述的概念。"可达性(reachability)"指的是,对于所有人来说,他们在多大程度上通过有限个步骤相互联络在一起:例如,谣言、观念或者资源在网络中传递的容易程度有多大。除了这些概念之外,巴恩斯(Barnes,1969)又加入了"派系"和"聚类(clusters)"这两个概念,用它们来区分网络中的社会群体。但是,这些概念并没有出现在巴恩斯(Barnes,1969)主编的经验研究文集之中。

人际关系是在由制度化的角色和地位构成的框架中得到构建的,而角色和地位存在于人际关系网之中,并要经过人际关系的再生产而存在。但是制度化的角色结构本身是网络分析的一部分,还是区别于网络分析,米切尔和曼彻斯特传统对此含糊其词。尽管某些曼彻斯特学者把制度化的角色结构看成是伴随人际关系网而存在的一种关系网,米切尔却常常把人际**关系网**与制度关系的**结构**区分开来。因此,米切尔倾向于对如下这种社会网络的"残留的"定义进行讨论:网络分析仅仅涉及到人际关系域,该域是在抽取出正式的经济、政治和其他角色之后剩下的领域(Whitten and Wolfe,1973)。这一点对社会网络分析后来在英国的发展具有重要意义。米切尔把社会网络分析看成是分析人际关系的一种**特殊**方法,就此而言,米切尔是不同于纳德尔的,因为后者热衷于建立结构社会学的一般框架,并且这种框架根植于形式网络分析。这一点命中注定了社会网络分析在英国的发展,因为随后这种研究不能吸引社区研究领域之外的追随者的注意力。

哈佛的突破

米切尔、巴恩斯和博特的观点在英国有深远的影响(参见 Frankenberg, 1966),他们的成功恰恰意味着社会网络分析与曼彻斯特人类学家的一些特定观念是一致的。也就是说,网络分析被看成是尤其关注"社群"类型的非正式人际关系,该方法也被认为是尤其关注个体网研究。其结果是,针对所有社会生活领域中社会网络的整体性质进行的开创性研究就不是在英国出现的。

实际上,社会网络分析的重大突破出现在哈佛大学。在霍曼斯出版其开创性论著十年之后,哈里森·怀特和他的同事们开始发表一些论文,从而大大地推进了网络分析。不久之后,又涌现出由这些论文作者的学生和同事撰写的大量论文,从而牢固地确立了社会网络分析的地位,使之真正成为一种结构分析的方法。

这次突破的关键因素在于数学上的两项创新(参见 Berkowitz, 1982 的有关讨论)。首先是群体代数模型的发展,该模型利用集合论对列维-斯特劳斯(Levi-Strauss)意义下的亲属关系和其他关系进行模型化分析。这导致了重新审视早期的图论及其他数学分支学科的研究,并试图利用代数模型对社会结构中的"角色"概念进行概念化分析(Boyd, 1969; Lorrain and White, 1971; White, 1963)。可以认为,怀特对"块模型(block modelling)"(参见本书第 7 章)的持续探讨推进了纳德尔早已强调的角色结构研究。第二个创新是关于多维量表的发展,这种"量表"技术可以把关系翻译为社会"距离",并在一个社会空间中绘制出这些关系。这些创新深深地扎根于卢因的场论研究传统,事实表明,它们是非常有效的分析方法(至于这些方法在社会学中的应用,参见 Laumann, 1966; Levine, 1972)。

上述两项创新结合在一起,推动以怀特为核心的新哈佛小组作出一些有影响的重要成果(参见 Mullins, 1973)。怀特是从芝加哥到哈佛的,他的研究工作与戴维斯及其他学者的研究有重要的关联,这些学者在整个 20 世纪 60 年代探讨了社会计量学的一些基本观点。哈佛小组主要进行具有数学导向的结构分析,对所有类型的社会结构进行模型化处理。在他们的研究中并不存在单一的理论关注点,他们的一致观点是利用代数概念对深层和表层的结构关系进行模型化分析。正是作为一种方法的网络分析才把他们团结在一起。格拉诺维特(Granovetter, 1973)的文章受到广泛接受,这也使这种观点在

美国社会学界大受欢迎,同时也激发了很多其他相关研究。尽管格拉诺维特的论文不涉及高度的数学技巧——或许正因为如此——该文使网络分析的理论和方法广为人知,进而在学术界具有合法的地位,格拉诺维特的工作功不可没(也可参见 Granovetter,1982)。尽管许多研究者在诸如共同体结构的分析领域等方面继续耕耘,其他学者却对诸如连锁董事会等现象感兴趣,从而把网络分析的关注点从纯粹的人际关系领域中转移出来。在进行这种研究的同时,他们也刺激了这些技术在其他的大量实质性领域的应用。哈佛小组的大多数努力——不再完全以哈佛为基础——都集中于在多伦多建立的国际社会网络分析网(International Network for Social Network Analysis,缩写为 INSNA),该网络受怀特的弟子韦尔曼和伯科威茨(Berkowitz)的领导,在社会网络分析的发展中占据核心地位[22]。

格拉诺维特和李分别作出的经典研究产生于哈佛学派最早的讨论。尽管二者的研究并不是明显的代数性质的研究,但是对于他人来说,这两项研究已经成为重要的典范。这至少不是因为它们与早期的社会计量学研究保持一种实质性的、分析的连续性。

在《谋职》(*getting a job*,1974)一书中,格拉诺维特首先对劳动经济学家在解释人们如何找工作方面的努力进行了批判式的考察。特别是,他希望探讨人们通过哪些方式利用其非正式的社会关系获得关于工作机会方面的信息。他感兴趣的是在信息传递方面的关系类型,这些关系"强"还是"弱",它们是如何随着时间的推移而维持的。为了达到此目的,他在波士顿郊区选择了一个男性样本,该样本中的人在过去 5 年内都有过跳槽经历,其中包括专业人士、技术人员和管理人员。格拉诺维特发现,非正式的个人联络是个人发现职业机会的主要渠道:56% 的被调查者依赖于这种方式,这对于高收入者来说更是如此。这些结论并不是非常令人振奋的,广义地说与早期的研究是一致的,格拉诺维特为自己设定的任务是区分出那些提供信息之人以及在什么情况下信息得以传递。

格拉诺维特认为,在决定获取职业信息的方法方面,"理性的"选择并不重要。个人实际上并不比较附加在不同信息来源上的收益和费用,人们并不急于"找"工作。相反,只要联络人主动提供信息,那么信息的获得都是偶然的。最重要的信息提供者往往是在工作中或者与工作有关的接触者。他们很少是家庭成员或者朋友,他们往往拥有不同于被访问者的职业。一个人跳槽的概率取决于他所接触到的与他的职业不同之人占总接触人数的比例。

格拉诺维特利用一种信息扩散模型来解释这些发现。这需要假定那些掌握职业信息的人会在一定程度上把信息告诉与自己直接接触之人,这些人又将信息传递给他们的直接接触者,依此类推。假定信息在沿着后续的关系

链[23]传递过程中随着时间的推移而减弱,我们就可能通过追溯传递网,发现获得该信息的人数以及这些人在网络中的各种不同的位置。因此,信息的获得将首先依赖于那些拥有并传递信息之人的动机如何,其次依赖于一个人的接触者在全部信息流中所处的"战略"位置(Granovetter,1974:52)。

格拉诺维特正是在这种论证中提出了一个现已著名的论断:"弱关系的强度"。强关系的重要性早已被理解。与一个人密切接触之人(家庭成员、亲密朋友、同事等)拥有很多重叠的接触者。他们相互认识并在很多场合下互动,因此,他们拥有的关于工作的信息也趋于一致。其中任何一个人获得的信息也容易传遍全体。反之,他们一般不会为网络之外的人提供什么新信息。他们获得的信息容易变得"陈旧",是其他人已经获得的信息。正是通过不经常接触的一些弱关系以及处于不同工作环境的人,才可能提供新的不同的信息。这意味着"熟人要比挚友更易于传递职业信息"(Granovetter,1974:54)。在格拉诺维特研究的几乎所有案例中,信息总是直接来自雇员或者雇员的一位直接接触者——一般来讲最多只存在一个中间人。通过多于两个中间人的关联是很少的。在获取有用的职业信息方面,短的弱关系链居于最重要地位。

相比之下,李关于《搜寻为人堕胎者》(*The Search for an Abortionist*)的研究要稍早一些。李要发现的是,在一个禁止堕胎的地区,女人是如何获得堕胎机会信息的。从事堕胎生意的医生不能打广告,他们必须在旅馆而不是诊所里进行手术。因此,那些希望堕胎者必须试着从他们的可能有过堕胎经历的朋友、熟人那里获得有关信息,因为这些人容易拥有这种信息,或者可以帮助联系到能够提供帮助之人。

为了研究这个过程,李与为人堕胎者和近期有过堕胎经历的妇女进行了接触。有意思的是,在构建样本的时候,她所采用的信息收集技术与那些妇女们使用的方法类似。与格拉诺维特类似,她混合利用访问法和问卷法来收集数据。在探讨了她们的生活中的各个方面,以及她们对待堕胎的态度之后,李转而考察她们如何找到替人堕胎者。这个过程涉及到要进行有根据的猜测,猜测谁会提供帮助,如谁可以提供替人堕胎者的姓名,或者进一步告知其他可能提供帮助之人。李发现,在成功地接触到一个替人堕胎者之前,一个女人平均要接触5.8个人,实际的接触者最少为1个,最多为31个。当然,许多接触是"行不通的","成功的接触链"长短不一,从一步到七步不等,其平均长度为2.8步。超过四分之三的成功接触链中包含两个或更少的中介人(Lee,1969:第5章)。接触者一般不是亲属或者掌权人士(如雇主、教师等),最重要的接触渠道是同龄女性朋友。

格拉诺维特和李都用简单的频次列联表来考察网络过程,仅仅对他们所

发现的网络关系的结构进行定性评论。确实,正如李所说的那样,在大尺度系统中极难找出各种重叠的个人网络的结构。无论如何,他们的研究工作对于社会网络分析在系统化分析方面的发展来说都极端重要。这些研究显示出,甚至最基本的社会网络方法都有重要力量,也极大地推动他们的哈佛同事们提出更精致的网络分析技术。

显然,作为一种具有引导性的概念和一类特殊的方法,社会网络分析的力量是不言而喻的。但是,把形式化的数学观念应用于研究社会网络,这种做法激励一些学者,使他们认为社会网络分析为新的社会结构理论提供基础。例如,巴恩斯和哈拉里(Barnes and Harary,1983)认为,我们可能从形式**概念**的利用推进到形式**理论**的应用。他们指出,只有当研究者超越出于纯描述目的的形式概念的利用(也可参见 Granovetter,1979),社会网络分析的预言才可能实现。数学是由一些定理构成的,这些定理指定了形式概念之间的决定性的逻辑关联。巴恩斯和哈拉里论证道,如果形式概念被证明是用来组织关系数据的有用的方式,那么这些定理也可用于那些数据。这样看来,来自形式数学的一些定理的应用"揭示了模型在真实世界中的含义,而这种含义可能是模型的设计者们都没有注意到或者利用过的。"(Barnes and Harary,1983)

有学者进一步指出,社会网络分析的发展已经指明了通向社会学理论新框架的道路,或者指明了重新坚持早期理论的途径。例如,其中尤其有影响者是对社会网络中的交换论视角的坚持(Cook,1977,1982;Emerson,1962,1964;Cook and Whitmeyer,1992),这一点也与广义的"交易论"研究(Bailey,1969;Boissevain,1974)和理性选择理论(Lin,1982,也可参见 Banck,1973 和 van Poucke,1979 中的讨论)有关。最近,有学者(Emirbayer,1997;Emirbayer and Goodman,1994)把社会网络分析确立为"关系社会学"的基础,从而可以取代那种强调文化和意义的研究,但是这种探讨与交换论或者理性选择理论无关[24]。

从长远的角度看,本书关注的要点并非社会网络分析法是否将在理论框架上体现出某种优势。毫无疑义,社会网络分析体现了一种面向社会世界结构的特定的理论取向。正因为如此,它必定与行动结构理论密切相关。但这并不意味着社会网络分析的最重要组成部分是任何一种特定的具体理论。本书将要详加阐述的是,社会网络分析是一种面向自身且有着一组特定的方法的社会世界的理论取向,但社会网络分析并非一种正式或具体的社会理论。

注 释

1 莫雷诺在 1933 年的一次讨论会上首次使用了社群图。1933 年 4 月 3 日的《纽约时报》(*New York Times*)报道了这件事。

2 平衡论以如下假设为基础,即个体将发现不平衡是令人不快的,因此将采取行动来建立或重建某种平衡。当然,这个心理学假设是有争议的。图论本身是一个纯粹的数学架构,它的运用也不需要这一心理学假设。

3 卡特赖特和哈拉里(Cartwright and Harary,1956)的书中概述的"位置"比这个更复杂,因为它们不仅根据三人组来界定平衡,还根据任一由线构成的"环"来定义平衡。关于环的观点在下面的第 6 章中将有进一步探究。事实上,戴维斯(Davis,1967)的后续工作表明,平衡性只能通过对三人组的考察来评估,他认为这种三人组相对于齐美尔认识到的那类三人组来说具有结构上的显著性。

4 可以用这个事实支持如下观点,即所有的凝聚力群体都通过把它们自身与一个陌生人外群体或局外人进行对比来确定它们的身份,不管这种群体是真实的,还是想象的。参见埃里克松(Erikson,1966)。

5 戴维斯误导性地将这些子群体说成是"派系"和"聚类",但是人们已经习惯性地定义这些术语,在此意义上说,它们既非派系也非聚类。出于这个原因,我用"子群体"这个一般性术语来概述他的立场。派系和聚类这两个术语将在第 6 章和第 7 章中进行界定。

6 梅奥自己对这项研究做出的影响深远的解释可以在梅奥(Mayo,1993,1945)中找到。

7 从我将在第 6 章的讨论中可明显看出,霍桑(Hawthorne)实验的研究者们使用的实质性概念"派系"与纯粹的根据社会计量学术语界定的概念是不同的,前者要比后者松散得多。

8 阿伦斯伯格和金博尔(Arensberg and Kimball,1940)关于爱尔兰乡村的研究也受哈佛大学的沃纳的指导,也与扬基城(Yankee City)研究高度相关。

9 如注释 5 的评论一样,这一观念并不符合社会计量学意义上的派系概念。沃纳的社会学观点也许可以适当地称之为一个社会"集合"或一个"人群"。

10 这不是一个纯粹的循环过程,因为可以从六类中的每一类中获取个体,这些个体可能是一个派系的成员,而该派系可能主要由第一类的成员构成。

11 他们从对这个堆叠矩阵的位置分析中着手建构像图——这还是一次在技术上的开创性的、初级的尝试,这一尝试也许再有 30 年也不会被普遍接受。有关这一方法的发展将在下面的第七章中讨论。

12 霍曼斯用"秩序"这一词来代替"方向"这个更有意义的术语。

13 事实上,他们对 18 位女性的分析仅仅是他们在旧城(Old City)中大约 60 个派系数据中进行的说明性选择。

14 霍曼斯在哈佛出现的时候，许多最初的研究者也在哈佛工作，这意味着这一声明是有一定证据的。然而，从洛特里斯伯格和迪克森（Roethlisberger and Dickson，1939）提交的实际报告来看，这一点还有争议。

15 虽然他反对帕森斯的理论立场，但“内部”和“外部”之分非常类似于帕森斯所使用的区分，这成为外部的“A”和“G”与内部的“I”和“L”之分的基础。正如我在下面的注释中将提到的那样，帕森斯是从罗伯特·贝尔斯（Robert Bales）（Bales，1950）有关小群体的研究的反思中提取他的概念的。

16 霍曼斯通过引入有关规范、地位和领导权的假设来进一步构建这个模型。其中一些模型是通过案例研究来说明的，但没有一个是专门的社会计量学研究。

17 博特不是该系的一个成员，但与其成员的关系紧密。

18 “人际秩序”也许是一个更好的术语，因为“人际网络”这个词语非常容易与“个体中心网”相混淆。后者在下文和第四章中有讨论。

19 这个区分与帕森斯以及在1950年代和1960年代的大多数社会学著述中做出的区分一样。参见洛克伍德（Lockwood，1956）对帕森斯的讨论中涉及的相关观点。这一区别也被哈贝马斯（Habermas，1968）再次发现，它也与霍曼斯关于“内部”和“外部”系统之间的区别有关。

20 这种关于多元性和关系数据的“堆叠”或联合的观点曾经是米切尔的数学关注的核心。

21 米切尔也简单地提到了关系的“频率”，但是最好把它看作关于行动（而不是关系本身）的一种测度。

22 INSNA是一个互相交换信息和彼此间提供智力支持的国际小组。该小组出版《联络》（*Connections*）这本研究通讯杂志，并主办了《社会网络》（*Social Networks*）这本杂志。它的网址是 http://www. heinz. cum. edu/project/INSNA①，该网址保留了《联络》的一些过刊内容，同时也提供了其他社会网络地址的链接。

23 这个衰减模型的基础是一个著名的观测，即在一排人中低声传递的信息最终将被扭曲。在格拉诺维特的模型中，流动的信息量将在链条的每一步中减少，因此，那些远离资源发送者的人就不太可能接收到有关工作机会的准确信息。

24 最好将关系式的或结构式的网络分析研究看作对文化研究的补充，而不是能够合并和取代它们的一种研究。参见怀特（White，1992a，1993）以及怀特和布伦特之间的讨论（Brint，1992；White，1992b）。

① 该网址现在已改为 http://www. insna. org——译注。

关系数据的处理 3

社会网络分析是作为一系列分析社会结构的方法而出现的,这些方法尤其允许对这些结构的关系方面进行考察。因此,这些方法的应用取决于关系数据而不是属性数据的可获得性。本章将探讨如何收集、储存这些关系数据以及如何准备进行网络分析。对于此类研究来说,在处理关系数据的时候需要考虑的一般性问题并没有什么特殊之处。在所有的社会科学数据中都会有如下问题:进入现场、设计问卷、抽取样本、处理无回答数据以及在计算机上存储数据等。这些问题在许多有关研究方法的一般性的和专门性的文献中已经得到了充分的探讨,这里无须重复。然而,在处理关系数据的时候确实又出现了大量特殊的问题。总的来说,由于这些问题在现存的研究方法方面的文献中没有涉及,因此,在进一步探讨社会网络分析本身的技术之前,很有必要首先对这些问题进行述评。

关系数据的整理

一旦收集到全部社会研究资料,必须用某种**数据矩阵**(data matrix)的形式加以保存(Galtung,1976),该矩阵是一种框架,其中的行或者经过编码的数据是根据一种高效或低效的方式组织在一起的。最简单的数据矩阵就是画在纸上的一个图表,这种图表包含很多行和列。当数据集合比较大,或者比较复杂的时候,就有必要在记录卡片或者电脑文件中存储数据矩阵。无论数据矩阵采取什么具体形式,其逻辑结构总是一个图表。在变量分析中,属性数据可以在一个案例-变量(case-by-variable)矩阵中加以组

织。每一个要研究的案例(例如回答者)用矩阵中的行来表示,列代表的是用来测量属性的变量。图 3.1 就是一个简单的数据矩阵图,其中包含了用来说明的变量。无论在纸上还是在电脑中,对于大多数标准的统计程序来说,数据都是以这种方式加以组织的。

		变量				
		年龄	性别	收入	……	等
个案	1					
	2					
	3			属性		
	4					
	⋮					
	等					

图 3.1 用来进行变量分析的数据矩阵

这种个案-变量数据矩阵不能用来表示关系数据。相反,关系数据必须从个案-隶属关系矩阵(case-by-affiliation matrix)这个角度来看待。图中的个案仍然是一些特定的能动者,它们构成了分析单位,但是各个隶属项则是这些能动者所卷入的组织、事件或者活动等。如此看来,矩阵的各个列代表了各个隶属项,从而可以区分出哪些能动者参与了何种隶属项。从这种案例-隶属矩阵中,我们可以引出关于行动者之间的直接关系和间接关系的信息。例如,图 3.2 就展示一个简单的案例-隶属矩阵,它表达了 3 个人(标记为 1、2、3)参与 3 个事件(标记为 A、B、C)的情况。如果有某人参与了某一事件,则在矩阵的对应部分交叉处记为"1",否则记为"0"。由图可见,3 个人都参与了事件 A,都没有参与事件 B 和 C。这样,根据这个表画出来的社群图就表达了一个简单的、在个体之间存在相互联系的三人结构。可以把这个社群图解读为:每个人在某一特定事件过程中都与另外两个人相遇。

		隶属		
		A	B	C
个案	1	1	0	0
	2	1	0	0
	3	1	0	0

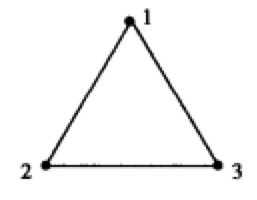
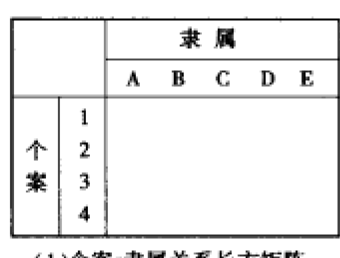


图 3.2 一个简单的矩阵和社群图

对于即使中等规模的数据集合来说,构建关于它们的社群图也是非常困难的。各条线将以各种角度相互交叉,从而形成了多种重叠关系,此时将丧失任何可见的、可以理解的结构。利用常规的手工绘画方法确实不可能绘出大型数据的社群图。因此,社会网络分析者早已试图寻找记录关系的替代方案。在遵循数据矩阵原则之基础上,人们接受的最主要方案是构建个案-个案矩阵,其中每个行动者都表达两次,一次表达在行中,一次在列中。各对行动者之间如果存在关联,则在交叉处用"1"表示,否则用"0"表示。这个观点可能不像社群图那样易于理解,因而需要费一些笔墨。

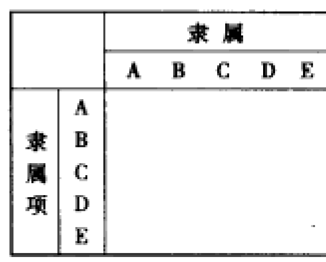
图3.3表达了社会网络数据矩阵的一般形式。对于初始的或者编码的数据来说,其最基本的形式就是我所说的个案-隶属关系矩阵,其中各行表达能动者,各列表示行动者隶属的项。这是一类2-模(2-mode)长方形矩阵,因为行和列表达的是不同的数据集合。因此,矩阵中的行数和列数一般不等[1]。从这种基本的长方形矩阵中可以推导出两个正方形的或者说1-模的数据矩阵。在个案-个案方阵中,行和列都代表个案,每个具体的矩阵格值展示了特定的一对行动者之间是否由于共同隶属于一个事项而关联在一起。因此,该矩阵展示了行动者之间实际存在的关系。就其所表达的信息来说,该矩阵与其对应的社群图是等同的。在第二类方阵中,行和列都代表隶属项,每个具体的矩阵格值展示了特定的一对隶属项之间是否由于拥有共同的行动者而关联在一起。在社会网络分析中,这种隶属-隶属关系方阵是尤其重要的,因为它通常可以阐明仅从个案-个案方阵中所不能明显看出来的社会结构的重要方面。



(1)个案-隶属关系长方矩阵



(2)个案-个案方阵



(3)隶属-隶属方阵

图3.3 几类社会网络矩阵

因此,一个简单的2-模长方数据矩阵可以转化为两个1-模方阵[2]。其中之一描述了初始阵的各行,另外一个则描述了初始阵的各列。我们并没有向初始阵中加入什么,这两个矩阵的生成无非是一种简单的转换罢了。长方阵和两个正方阵都是针对同类关系数据的相同表达方式。

在社会网络分析中，长方阵一般被称为"'发生阵'('incidence' matrix)"，而两个正方阵又称为"邻接阵(adjacency matrices)"。这些术语来源于图论，在后面的章节中我们会做出具体说明。在这里，我们仅仅记住这些术语就可以了，因为它们是在关系数据矩阵中最常利用的术语。大多数网络分析技术都涉及对邻接阵的直接操作，即涉及先把初始的发生阵转换到两个邻接阵。因此，研究者必须理解自己所使用的数据的形式(不管它是发生阵还是邻接阵)，并且理解在网络分析程序中所隐含的一些假设，这是尤其重要的。

研究者收集到关于一些个案及其隶属关系的二模数据的时候，一般来讲，最好把这些信息整理成一个发生阵，然后从中构建出在社会网络分析中使用的邻接矩阵。然而，在某些情况下，研究者可能直接利用个案-个案形式来收集关系数据。例如，在一个小群体中的择友情况即如此。在这种所谓的直接社会计量选择数据的情形中，可以立即在邻接矩阵中把信息组织起来。用不着复杂的讨论即可知道，在这种情况下不存在与该矩阵对应的发生阵，也不存在与之互补的邻接隶属矩阵。当然，其原因在于全体行动者都仅仅共享一种隶属性，这就是如下事实——即他们已经相互选择对方为朋友[3]。

出于许多社会网络的目的，"个案"和"隶属"之间的区别也可能是人为设定的。例如，在对 18 位妇女参与 14 个社会事件的研究中，一般要把妇女看成是个案，把事件看成"隶属项"，这才是可以理解的。这确实与标准社会调查中把行动者看成是案例这一点是一致的。但是，对于诸如"多个组织的成员如何共享"这样的现象来说，就很难明确哪个是"个案"，哪个是"隶属项"。此类研究关注的是，各个组织的成员在多大程度上有重叠，它们招募新成员的方式在多大程度上相似。在社会学意义上，"组织"及其"成员"都是能动者，因此拥有相等的被看成是"个案"的权利。成员可能被看成是个案，在这种情况下，他们所在的组织就被看成是他们的"隶属项"了；或者相反，把一些组织看成是个案，把这些组织共享的成员看成是隶属项。出于进行网络分析之需要，究竟把哪些行动者看成是个案，这仅仅依赖于哪类行动者在研究设计中被认为是最重要的。

这种决定已经体现在先前的有关抽样的决策之中。如果认为组织具有极端重要的意义，就可以选取由多个组织构成的一个样本进行研究，并且在后续的分析中出现的人物只能是那些恰好为这些组织的共同成员的人物。在这种研究设计之中，各个组织拥有一种理论上的优先性，并且一种比较合理的想法是认为一些成员表达了组织之间的隶属关系。然而，就网络分析技术而言，把何者看成是"个案"已无关紧要。因为不管做出怎样

的选择，分析程序都是一样的，并且研究者的任务就是决定针对哪一项可以进行有意义的社会学解释[4]。

因此，一般情况下，可以把“个案”和“隶属项”之间的区别看成是网络分析研究设计中的一种纯粹约定性特征。这种约定性的深层次方面是把个案置于发生阵的行位置上，而“隶属项”放在列位置。这要基于在属性分析中所作的一些约定，因为在属性分析中，个案一般用行表示，变量用列表示。

如果数据矩阵被用作关系数据的一种基本组建框架，那么我们还必须理解另外一些约定。这些额外的约定被认为是网络分析中最佳做法之基础，因为它们有助于在研究讨论中保持最大的明确性。大多数读者可能对基础数学中各种约定的重要性比较熟悉。例如，在画常规变量图的时候，一般约定垂直轴为因变量，并记为 y 轴；水平轴用来表达自变量，并记为 x 轴。这种约定会防止我们在阅读图的时候出现混乱，并且保证了关于图的任何陈述都不是模棱两可的。与之类似，关于关系矩阵的一些约定也出于同样的目的。

在讨论矩阵的时候，我们约定把矩阵的行数记为“m”，列数记为“n”，并且在描述矩阵规模的时候，常常先指出行数，后指出列数。用这样的方法，我们可以把一个矩阵的大小概括为一个 $m \times n$。例如，图 3.3 中的发生阵就是一个 4 × 5 矩阵。在描述任何一个特定格值的内容的时候，也常常先给出该格的行，然后给出列，并且用字母“a”来代表该格包含的真实值。这样，在第 3 行、第 2 列交叉处的矩阵格值就记为 $a(3,2)$。推而广之，我们约定行用 i 表示，列用 j 来表达。因此，$a(i,j)$ 指的就是第 i 行、第 j 列交叉处的值。也就是说，一个格值的一般表述方式为 $a(i,j)$，研究者可能继续给出 i 和 j 的具体取值。这些约定在图 3.4 中进行了总结。

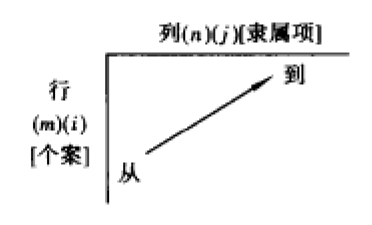
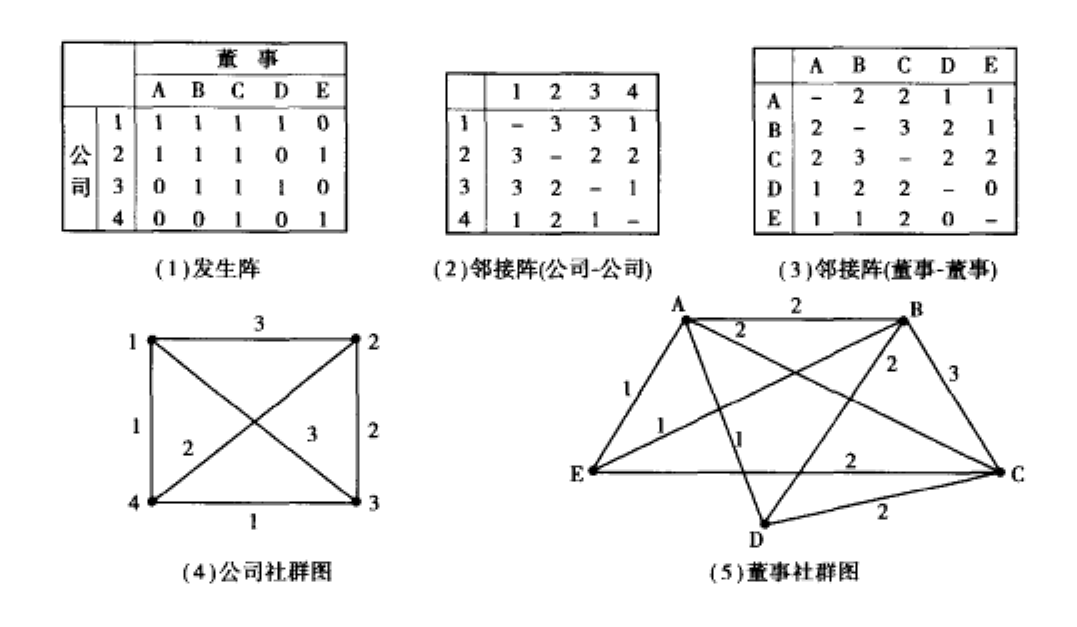


图 3.4 矩阵中的约定：最佳做法

下面用一个具体例子来说明如何应用矩阵来分析关系数据。图 3.5 包含了一个虚构的各个公司之间的连锁董事会数据。当某个人兼任两个或者多个公司董事的时候，我们就说存在一个连锁董事。

他（她）出现在两个董事会之中，在两个公司之间建立了联络。在关于连锁董事会的很多研究中，人们的核心关注点是公司。因此，他们一般把公司看成是个案，公司因而出现在图 3.5 的发生阵的行中。该矩阵的各个列即为隶属项，是各个公司共享的董事。矩阵的每个格值用二进制数“1”



		董事				
		A	B	C	D	E
公司	1	1	1	1	1	0
	2	1	1	1	0	1
	3	0	1	1	1	0
	4	0	0	1	0	1

(1)发生阵

	1	2	3	4
1	-	3	3	1
2	3	-	2	2
3	3	2	-	1
4	1	2	1	-

(2)邻接阵(公司-公司)

	A	B	C	D	E
A	-	2	2	1	1
B	2	-	3	2	1
C	2	3	-	2	2
D	1	2	2	-	0
E	1	1	2	0	-

(3)邻接阵(董事-董事)

(4)公司社群图

(5)董事社群图

图 3.5　连锁董事关系矩阵

或者“0”来表示,它们分别表达了每个董事是否出现在每个公司中。由图可见,公司 1 有 4 个董事(A、B、C、D),而董事 A 同时也是公司 2 的董事。这意味着公司 1 和 2 之间有关联。邻接阵(2)表达了所有公司之间的连锁董事情况。在这个矩阵中,每个格值不仅表示连锁董事是否存在,还表达了每对公司之间共享董事的个数。每个格值包含的是实际数,而不仅仅是简单的二进制数,因为公司之间共享的董事可能多于一个。可见,公司 1 和 4 只有一个共同董事,即 C。而公司 2 和 3 则有两个共同的董事,即 B 和 C。这一点可在初始发生阵的列中得到确认,从中可见,董事 C 身处公司 1 和 2 的位置上,董事 B 和 C 每个人都坐在公司 2 和 3 的位置上。

从对邻接阵的这种最简单的分析中可以看出,关系的强度可以通过关系本身包含的“连锁”数来测量。那么,“最强的”关系存在于公司 1 和 2 之间以及公司 1 和 3 之间,每组都包含 3 个董事。“最弱的”关系指那些只共享一个董事的公司之间的关系。公司社群图(4)清楚地表明了矩阵的结构,每条线上的数字表示线的强度或取值。这种社群图也可以用其他方法画出来:例如,用线的粗细表达关系的强弱,或者用一条、两条或三条平行线连接的点代表关系强度。每种方法表达的关于矩阵结构的信息都相同。

前文已经指出,从一个发生阵中可以引申出两个邻接阵。在本例中,不但可以从发生阵中引出公司-公司邻接阵,还可以引出董事-董事邻接阵。在图 3.5 中,该矩阵及与之相关的董事社群图表明了各个董事之间存在的

关系,即他们担任了同一个公司的董事。例如,B 和 C 之间具有强关系,他们在三个独立的公司董事会(公司 1、2 和 3 的董事会)上相互谋面。A 和 D、A 和 E 以及 B 和 E 之间的关系相对较弱,因为它们之间仅仅具有一对董事关系。董事社群图也表明了如下所示的一些社群图观念,如 D 和 E 相对其他董事来说居于网络的边缘地位:它们的关联较少,关系一般较弱,实际上它们相互之间没有联系。

图 3.5 中的两个邻接阵也表达了社会网络分析中的更一般性的思考。首先,很重要的一点是注意矩阵中从左上角到右下角的对角线。在矩阵分析中,这种特殊的对角线仅仅指的是"对角线",因为对角线上的值不同于矩阵中的其他值。在一个方阵中,对角线的值表达了任何个案与自己的"关联"。在某些情况下,这种关联没有实际意义,但有时候它可能是网络的一个重要特点。例如,在图 3.5 的矩阵(2)的对角线的值就代表每个公司与自己的关系。在本例中,这没有什么特殊的意义。一个公司通过其所有董事而建立联系,这一点确实为真,但无意义,因为我们关注的是公司之间的关系。因此,对角线上的值是不存在的,应该在分析的时候忽略掉。在网络分析中,如果存在模棱两可的话,那么许多技术程序要求研究者特别指出对角线的取值是否包含在分析范围之内。因此,研究者必须清醒地认识到矩阵中的对角线的地位,并且要理解某些特定的程序是如何处理对角线上的值的。

图 3.5 也表明邻接阵关于其对角线"**对称**":每个矩阵的上半部分是下半部分的"映射"。其原因在于这种数据所描述的是"无方向的"网络,例如,公司 1 对公司 2 的关系等同于公司 2 对公司 1 的关系。二者之间关系的存在被认为是独立于如下问题的,即关系是否涉及权力和影响在一个方向(而不是另一个方向)的运作。因此,在一个无向网络邻接阵中,全部关系信息都包含在矩阵下半部分之中;严格地说,上半部分是多余的。因此,许多网络分析程序仅要求知道邻接阵的下半部分,不要求有全部矩阵。对于无向网络来说,这种分析方法没有损失任何信息。

"无向"数据是最简单,也最容易处理的一类关系数据,也许现在有必要花点时间讨论一些比较复杂的数据类型。在变量分析中,最重要的一个问题是适用于一个变量的"测量层次"是什么。也就是说,属性数据应该被测量为定类变量,还是定序、定距、定比变量。只有选择测量层次之后,才可以据此决定选择适用于该数据的特定分析程序。与之类似,根据关系"有""无"方向以及/或者是否"多值",关系数据也存在同样的"测量问题"。图 3.6 根据这两个维度对关系数据作了 4 个层次的测量。

最简单一类关系数据(类型 1)是既无方向的又二值的数据。图 3.5 中

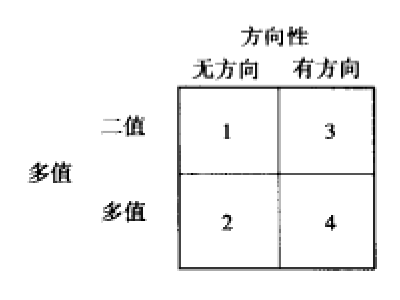


图 3.6 关系数据的几种测量层次

的发生阵数据即为此类型，而其中两个邻接阵则包含第二类关系数据：其关系是无方向的但却是多值的[5]。前文已经指出，图 3.5 中邻接阵所包含的“多值”数据（类型 2）来源于初始发生阵的二值数据。多值主要表达了关系的强度而不仅仅是关系的有无。上一章讨论的与平衡理论相关联的符号数据（signed data）也是关系数据，其中的符号“+”或者“-”赋予每一条线。这些关系可以看成是二值关系和多值关系之间的一种中间类型。这种数据不仅表现了关系存在与否——因为存在的关系恰好是通过给出的正号和负号来表示；而且表达了关系的性质，这种性质仅仅是通过正负极，而不是通过实际取值来表达的。当然，可以把一个记号和一个取值结合在一起，从而赋予关系数据以各不相同的编码，例如从 -9 到 +9。在这个过程中，取值不能仅代表各个案例之间共同隶属的事项的数目，因为它们之间共享的隶属项数不可能是负数。因此，这些赋值必定是关于关系强度或者紧密性的其他测度。当然，这种过程要依赖于一种社会学论据，该论据应该提供一些稳固的理论和经验原因，从而说明为什么要以这种方式对待数据。

多值数据总可以转变为二元数据，其方法是把矩阵二元化，用一个临界点对各个格值进行“切分”，尽管这样做会损失一些信息。在切分过程中，研究者只选择那些高于某一层次值的关系，并视之为显著关系。高于该层次的所有值都被切开，从而构建一个新的矩阵，其中高于此层次的值都用“1”来表示，等于或低于该值的格值都用“0”来代替，这样就构造了一个新的矩阵。在网络分析中，这种切开数据矩阵的过程是非常重要的技术，第 5 章将对此详加讨论。有向数据也可以采用二值或多值的形式，也可以应用同样的切割程序把多值有向数据（类型 4）降低为二元有向数据（类型 3）。把有向数据下降为无向数据也可以，此时仅仅忽略关系的方向即可。此时研究者可能考虑的重要事情仅仅是关系是否存在，而无需关注其方向。在这种情况下需要忽略关系的方向，这是可以理解的。就此而言，应该进一步给出一个恰当的矩阵约定。在包含有向数据的邻接阵中，通常约定关系的方向是“从”行元素“走向”列元素的。如此看来，在一个有向阵中，格值(3,6)表达的意思是从个体 3 指向个体 6 的关系或者强度。而个体 6 到个体 3 的关系应该表达在格值(6,3)中。这个约定体现在图 3.4中。正因为一个有向阵一般不关于对角线对称，因而必须要通盘考

虑整个矩阵，而不仅仅关注其下半部分。

复杂类型的关系数据总可以化归为简单的数据形式，最后，**任何**形式的关系数据都可以看成是无向二元数据（类型1）。因此，在全部社会网络分析技术中，适用于分析此类数据的技术已经得到了最广泛的应用。但是，反向的操作是不可以的，即不能将简单数据转换为复杂数据，除非在初始数据矩阵之中还包含额外的信息[6]。

研究者必须非常谨慎地对待关系数据的本质。尤其是，他们必须确认数据的测量层次具有社会学意义。例如，在研究公司连锁关系的时候可以利用多值数据，这种研究依赖于一些假设，一些关于多个董事职位的重要性是否具有适当的假设。例如，研究者可能假设，两个公司之间共享董事的数量是二者之间关系**强度**或者**紧密性**的一个指标（indicator）。这样看来，有4个共同董事就意味着这两个公司之间的关系可能比仅有2个共同董事的两个公司之间的关系"紧密一些"。但是，这是一个有效的社会学假定吗？如果不是，就不能利用数学程序计算。数学本身不能为研究者提供答案。某些特定的数学概念和模型之间的相关性对于需要给出有充分根据的社会学判断的研究者来说总是一个问题。即使确定了可以应用多值数据，研究者也必须对在数学程序中可能包括的其他假设保持一种敏感性。例如，一个程序把取值看成是定序还是定比变量？在前一种情况下，取值为4仅仅意味着比取值为2强一点；在后一种情况下，取值为4的关系就意味着比取值为2的关系强**两倍**。再强调一遍，对测量层次的选择是一个社会学问题，而不是数学问题。

关系数据的存储

针对非常小的数据集合可以进行很直接的分析。例如，我们完全可以手工构建一个4人或者5人群体的邻接阵和社群图。然而，如果网络规模扩大，这种分析将变得非常复杂。当处理包含多于10个案例和5个隶属项的数据集合的时候，除了用计算机来分析外别无他法。计算机运行过程不但节省了大量时间（例如，霍曼斯在研究18个女性参与14个事件的过程中，利用手工方法对矩阵进行重排，而这一过程用电脑来做最多需要几秒钟就可以完成），它还可以做一些利用手工不可能做到的分析工作。

如果关系数据得到恰当的存储，就可对它进行高效的管理和操作。如果需要利用计算机进行网络分析，那么这意味着要考虑如何把数据矩阵的

逻辑结构转译为一个计算机文件,这一点是很重要的。第一步常常是对行动者或事件进行分类,以便产生各个名单,从而可根据名单分析它们之间的关系。例如,在研究连锁董事成员的时候,就涉及要产生身处各个目标公司中的一个董事名单,以字母顺序进行分类,然后区分出那些出现两次或者两次以上的名字。进行这种分析的最直接的方法就是利用文本编辑器或者字处理程序来生成一个数据文件,因为名字可以以文本的形式输入,并且可以进行分类和编辑。许多字处理程序都可以根据字母顺序或者数字顺序对数据进行分析和操作[7]。

对数据进行此类分析的最常见的结果是以"联系名单(linked list)"格式出现的数据。在联系名单中,文本中的每一条线表达一个案例及其各个隶属项。例如,它可能是这样的数据,即给出一个董事的名字,后面跟着的是该董事所在的全部公司的名字。然而,通常情况下,仅仅利用字处理程序不能把它转换为发生矩阵(如图 3.3 所示)。除非使用者希望进行较难的——同时也是易于出错的——手工操作,否则需要把数据直接转移到社会网络分析程序如 UCINET 之中。在这个程序中,联系名单数据可以输入进去,并且被内在地转换为一个发生阵。该程序还允许以它们的初始联系名单的形式,对新数据文件进行直接编辑。

在 UCINET 中,联系名单格式是以一个工作表(spreadsheet)的形式出现在屏幕上的,并且事实上,大量的数据处理任务都利用诸如 EXCEL 这样的工作表程序来完成。如果没有社会网络分析程序的话,那么一个工作表可以直接从一个字处理程序中输入联系名单。尽管工作表仍然被看成是财会人员和股票市场分析者广泛使用的财务工具,在本质上它乃是一个电子矩阵操作程序。即使最简单的工作表也可以用来储存和组织关系数据,可被用来在一些文件中准备这些数据,以便被许多其他更专业的程序读取。现在,一些工作表程序随处可见,因而值得把它们看成是对社会网络数据进行保存和操作的一种基本系统。如果数据已经从关联名单转换为二值或者多值形式的矩阵,那么可以用工作表来计算一些基本的统计测度,如行数和列数、频次分布、相关性等。其中的一些测度可以转换成屏幕上的图形,并且打印出来。例如,频次分布可以立即用一个直方图或条形图展示出来。尽管在许多工作表中内含的主要数学函数都是那些更适用于进行变量分析的财务和统计程序,也有很多工作表程序包含了一些矩阵数学计算项,从而可以用它们来计算各种不同的网络结构属性[8]。

在工作表中储存的数据可以很容易地操作,从而解决了数据准备中存在的实际问题,而这些问题曾经困扰过网络分析者。例如,基本上所有的工作表都允许对行和列按照字母顺序或者数量顺序对数据进行自动重排。

工作表的"范围(range)"选项可用来对矩阵的特殊部分进行分类,从而复制成一个新矩阵。例如,如果以文件形式存储了一个人们之间的朋友关系矩阵,就可以分别选择男性或者女性数据单独作分析。甚至可能把发生阵转换为与之对应的一些邻接阵。然而,这种工作表最好在得不到其他社会网络分析专业软件的时候才加以应用。应该主要用工作表来存储数据,并执行重排和操作等直接的数据管理功能[9]。

社会网络分析的两个最常用的软件——UCINET 和 STRUCTURE——都以简单的矩阵形式存储数据,很容易把工作表中的相应文件转换为上述两种文件形式[10]。在大多数情况下,最好尽早把文件转换为上述一种专业软件的文件形式,只有当属性数据必须添加上并且用来进行统计分析的时候,才把这些数据再读回到工作表中。事实上,在这些情况下,最好把数据输出到诸如 SPSS 这样的专业统计软件当中。

功能最强大的一种网络分析软件是 GRADAP,但是,这种软件所使用的数据形式不同于到目前为止我们讨论的矩阵结构形式。GRADAP 可以与 SPSS 交换数据,但是不能处理直接输入的发生阵和邻接阵。在一个工作表、字处理程序或者 SPSS 文本编辑器中可以产生一个 GRADAP 数据文件,但是这种数据的形式不像数据矩阵那样具有直观性。GRADAP 的目的是把关系数据彻底转换为图论术语,因此要求对构成数据的点和线有非常明确的界定。GRADAP 要求把邻接阵(而不是发生阵)本身分为两个独立的文件:"点文件"和"线文件",前者列举出各个案例,后者列举出每个关系。一条线用其两个端点来界定。当研究者拥有直接的社群图数据的时候,或者如果实际的关系模式可以根据一定的方式观察出来的话,这种界定就很少产生问题。但是,当数据是以关联列表或者发生阵的形式表达的时候,就很难生成必要的文件。即使在工作表或者数据库程序的帮助下,在产生 GRADAP 文件之前也需要进行大量的手工操作。

图 3.7 展示的是与图 3.5 中包含的发生阵数据对应的一个 GRADAP 线文件的形式。在该网络中包含 12 个连锁的董事,这个值可以从图 3.5 中邻接阵(2)下半部分值的总和得到确证。每一个连锁都被计算为一个独立的"线",从而作为数据输入 GRADAP 中,因此"线文件"中包含 12 项[11]。其中的每一项都用"头"点和"尾"点来区分,关于线的进一步信息(例如与之对应的董事的姓名等)可以加入文件之中[12]。通过对比图 3.5 和图 3.7 可以确认,GRADAP 线文件包含了发生阵中的全部信息,这也是为什么 GRADAP 能够根据线文件无形地构造出两个邻接矩阵的原因。

一旦创建了一个 GRADAP 文件的结构,该程序就提供了一些强大的数据管理工具,几乎与一个针对关系数据的工作表管理系统一样运作。然

线的代号	1	2	3	4	5	6	7	8	9	10	11	12
头	2	2	2	3	3	3	3	3	4	4	4	4
尾	1	1	1	2	2	1	1	1	1	3	2	2
线的信息	A	B	C	B	C	B	C	D	C	C	C	E

图 3.7 一个 GRADAP 线文件

而，即使进行最简单的分析，该程序也要求对图论知识有完全的了解。因此，对于初学者甚至有时候对于网络分析者来说，UCINET 加上字处理程序才能提供最佳的工具。对于那些能够使用其他程序，从而生成线文件的高级使用者来说，GRADAP 就有很多优点了。笔者将在本书附录中对这些程序稍加详细的讨论。

关系数据的选择

本章前两部分考察了关系数据的本质，并且指出如何对关系数据进行组织和管理，以便进行网络分析。在澄清了整理和存储数据的方式之后，现在才可能考察一些遗留的问题，即涉及数据收集过程的问题。我们已经指出，在这个领域中很少会产生特殊问题，但数据的选择问题却给社会网络分析带来一些值得考虑的难题。这些选择问题涉及社会关系的界限以及从样本中获得关系数据的可能性问题。

在研究小规模社会网络的时候，常见的策略是把一个群体的所有成员区分出来，追溯他们之间的各种各样的关系。但是，这远非可直接操作之事。各种社会关系都是一些社会建构，是在对群体成员所处情境进行定义的基础上产生的。例如，“亲密朋友”关系对于不同人来说可能意味着不同的事情，因为不同人对“亲密”的理解不同。如果仅仅要求回答者指出他们的“密友”，那么研究者不能保证所有的回答者都对“亲密性”有相同的理解。对“亲密性”有严格界定的回答者将给出围绕自己的比较窄的界限，而那些对“亲密性”有宽泛理解的回答者辨识出的边界将扩大。因此，亲密朋友群体的界限就因人而异。研究者根据这些人群确定出来的任何边界可能完全是人为的——只不过是提问的问法的结果罢了。另一方面，如果研究者明确地界定“亲密”（如根据互动频次来界定“亲密”），那么他（或者

她)就会把自己的这种亲密定义加到回答者的身上,此时朋友的界限仍然是人为的。

这个问题很重要,因为研究者常常坚持关于关系系统界限的一些不现实的观点(Laumann et al., 1989)。人们常常假设个体之间的社会关系只限于所研究的特定群体或者区域之中。就该区域之外的关系被忽略这一点来说,社会网络分析将不是对全部网络的完美表达。在非正式群体(如街头团伙)的案例中更为明显,因为该群体的界限是松散地界定的,团伙成员的活动也超出其核心成员之外(Yablonsky,1962)。对于比较正式的群体来说,情况也是如此。例如,在克尔和费希尔(Kerr and Fisher,1957)讨论的"工厂社会学(plant sociology)"中,他们关注独立于更广的经济之外的特定工作组和办公室的空间界限。这种研究使研究地点脱离它所嵌入的较大的区域、国家和国际系统。限定在局部工作情境的研究可能不会区分出那些超出工厂之外的关系。

斯泰西(Stacey,1969)也以同样的方式批判了这种局域研究,因为这种研究假定"社群性"的团结关系局限在当地的社会系统之中。她认为,必须把这些关系看成是与更大的经济、政治系统纠缠在一起的。同样,劳曼等人(Laumann et al.,1983:31)认为,对资金在网络中流动的局部性质的研究不应仅限于地理上的局域内部。在资金流通过程中,许多最重要的机构恰恰位于该地理区域之外,如一些州政府机构、地方性银行和国家银行等。如其所述,如果这些机构对于资金的流动来说要比一些地方性的组织和机构重要的话,那么一种建立在地方基础上的研究计划就可能面临如下这种可能性,即它对相关的交易网络结构方面的看法有可能是非常不充分的。

所有这些问题都指向如下事实,即网络界限的确定不仅仅是一个对所研究的情景区分出自然的或者明显的界限的问题。尽管"自然"的界限确实可能存在,但是,在一项研究计划中,界限的确定是一项在理论上有充分证据支持的如下决策的产物,即要确定在所研究的情景中什么是重要的。例如,在研究政治关系的时候,必须认识到什么是"政治的",它如何与"经济的"、"宗教的"以及其他社会关系区分开来,并且相关政治单位的界限确定等都要有理论根据。研究者要进入一种概念考证和模型建构的过程,而不仅仅是一个收集"事先形成的"数据的过程。

假设界限已经确定,下一步任务就是确定目标总体。对此任务的研究通常有两类:"位置"研究和"声望"研究[13]。在**位置研究**(positional approach)中,研究者首先从得到正式界定的特定位置的占据者或者群体成员中进行抽样。首先,确定感兴趣的群体或者位置,然后从中抽取出位置的占据者或者成员。除非所研究的总体特别小,否则都要求有某种列举出

来的列表，该表要涵盖整个目标群体。此类研究策略的例子有来自诸如一个学校班级、一个村庄、一个工作小组，或者如政治精英、公司职位等体制的样本。对于位置研究来说，一个我们熟悉的问题是如何决定应该包括哪些位置。例如，精英研究常常受到批判，因为这种研究只区分出机构等级中的高层位置，特别是当研究者没有一个临界标准，从而把机构等级中的**“高层”**位置与其他位置区分开来的时候，此类研究就更受到批判。当然，这个问题也反映了上述讨论的一般性的界限问题，重要的是，一些特定位置是包含在内还是排除在外，对此研究者要在理论和经验上拥有充分可辩护的原因。

有时候这需要有一个假设，即认为总体内部“自然”会有一些子群体。例如，在对商业连锁关系的研究中，人们常常关注存在于一个经济体系中的“250 强”公司[14]。这种研究设计就涉及如下假设，即第 250 个和第 251 个公司之间的差距自然构成了大公司和中等规模公司之间的自然边界。然而，这些边界很少能够得到精确的界定。规模从大到小有一个持续的衰减过程，并且尽管可能区分出一些斜率改变的点，一般也不可能给出明确的界限来。确实，大多数此类研究都没有考虑到斜率变化的总规模分布，而只是利用一种人为的，**先验的**临界标准：尽管有人研究了 250 强公司，其他人却研究分布中的前 50 强、前 100 强或者前 500 强的公司[15]。

在位置研究中，有时候研究案例的选择可能遵循早期的选择隶属项之时的决定。例如，一个董事职位可以看成是一个人与一个公司之间的隶属关系，而研究者可能已经决定把关注点限定在特定的一些公司上。在这种情况下，选择什么样的董事来研究，这就取决于在选择公司的时候所坚持的标准[16]。

当不存在相应的位置，也没有全面的名单，或者能动者自己的知识对于总体边界的确定来说非常重要，在这些情况下，就可以进行**声望研究**。在此类研究中，研究者要考察由一些信息灵通的线人提名的所有人或者部分人。列表中包括的人名是那些被提名成为目标总体的成员的名字。例如，根据不同的研究目的，线人可能就“哪些人是社区中有影响力的成员”、“商业中名望高的人”等进行提名，然后把这些提名组合成为研究总体。显然，在声望研究中，线人的选择是极其重要的。研究者必须有充分的理由相信线人对目标总体有很好的认识，并且能够准确地汇报出来。至于情况是否如此，常常只有当探究结束之后才知道，因此在研究策略中存在一定的循环往复性。因此，在选择线人的时候，研究者应该努力诉求理论上和经验上的理由，尽可能使线人独立于需要加以研究的特定社会关系。

这种要求并非总是可能的，因为在这种声望研究的一种特殊变体，即

应用所谓的"滚雪球"技术的研究中,所遵循的程序恰恰相反。在这种研究中,先研究一小部分线人,再请每个线人提名其他人。再对这些被提名者进行访谈,并要求进行进一步提名。随着提名的进行,就像一个雪球那样建立了一群被访谈者。最后,在每一轮访谈中将很少有添加的提名者。在滚雪球方法中,社会关系本身被用作建构群体的一种关系链。然而,就其本质来说,一个滚雪球样本一般是围绕着作为开始点的一个特殊个体的关系而组织的。因此,所选择的方法趋于决定所得到的社会网络的许多关系属性。这种网络是根据一个相互联系的行动者群体中的各种关系建构起来的,并且如劳曼等学者所评论的那样,"由滚雪球抽样方法构造出来的一个网络是关联紧密的,这一点几乎不用言说。"(Laumann et al. ,1983:22)

当发生阵中的各列是真正的隶属项,并且研究者的目的是把它们从案例中单独选择出来的时候,就需要最后一种选择性的策略,该策略既不是位置研究也不是声望研究。也就是说,这种研究关注于对人们所参与的活动和事件的选择,这种选择与任何用来区分人们的位置或者组织是无关的。例如,在对纽黑文市(New Haven)的研究中,达尔(Dahl,1961)应用"参与关键决策的制定"作为选择的基础。因此,"参与一项决策"就被看成是一个"隶属项",据此,人们可以被赋予一个二值的或多值的数值,这些值要独立于他们在组织中占据的任何位置。这允许达尔(他可能也认为)对不同类别行动者的相对权力进行评价,而不用假定权力是位置的一种自动关联项。与之类似的另一项研究是由戴维斯(Davis,1941)及其同事在美国东南部地位(*Deep South*)中做出来的,他们研究了一些社会事件,最终用一个矩阵来展示 18 位女性参与 14 个事件的情况。当然,在此类研究策略中存在的问题是如何对各个隶属项的选择进行论证:最重要的事件是否被选出来了?什么是一个"关键问题"?因此,选择真正的隶属项恰恰与在选择案例过程中涉及的问题是相同的。事件和隶属项之所以被选出来,是因为人们认为它们具有客观上的显著性(这是位置研究的一个变体),或者因为一些博闻之线人认为这些是重要的(这是声望研究的一个变体)。

到目前为止,笔者探讨的主要是根据完全的或者准-列举(quasi-enumeration)的方式来选择整个总体的问题。但是有时候必须利用样本资料来分析,此时这些问题就变得复杂了。在小群体研究中很少出现抽样问题,因为这时候我们完全可以列举出全体的群体成员,列举出他们两两之间的关系。然而,当研究大尺度社会系统的时候,难以做到完全列举,此时将特别出现一些难以解决的抽样问题。所需要的资源提供者的规模常常会阻止对大总体进行完全的列举,但是,即使这种研究是可能的——例如在对一个总体进行普查的时候——作为结果的数据矩阵的规模将使得任

何分析成为不可能。在进行网络分析之前，一般首先要构造正方形的邻接矩阵，矩阵数据可能规模很大。例如，一份拥有 1 000 个案例和 50 个变量的属性数据要在矩阵中包含 50 000 项数字。计算技术的进展已经使得对这些矩阵的统计分析变得容易一些。然而，就关系数据来说，由 1 000 个案例构成的案例-案例矩阵将包含1 000 000项数字。在一个拥有 5 000 人的相对较小的村落中，邻接矩阵将包含 25 000 000 项，这已经超出了大多数现有的计算机和软件的分析能力[17]。对于一个拥有几百万居民的国家来说，数据之巨大几乎难以想象，而用来分析它的计算机能力在科幻小说领域之外是不存在的。

当然，在前计算机时代，同样是这些问题却导致了抽样技术的发展，从而允许使用诸如容量为 1 000 的样本，而用不着对一个拥有成千上万个体的总体进行全部的列举。关于抽样的统计学理论一般要确定一些条件，在这些条件下，可以分析从一个案例样本中收集到的属性数据，进而推广到对大总体的估计。因此，可以假定，来自大总体的抽样可以为社会网络分析提供同样的可解决方案。图 3.8 给出了社会网络分析中的一种理想的抽样过程的概要式说明。在构成一个整体网的各种类型社会关系的复杂系统中，将涉及一个特定的能动者群体。在该关系系统中，社会学家们可能把“局部网”确定为那些构成了经济关系、政治关系、宗教关系等的网络。当坚持一种完全列举策略的时候，研究者试图保证在全部相关关系上获得全部的信息，因而可以构建一些充分的局部网络模型。

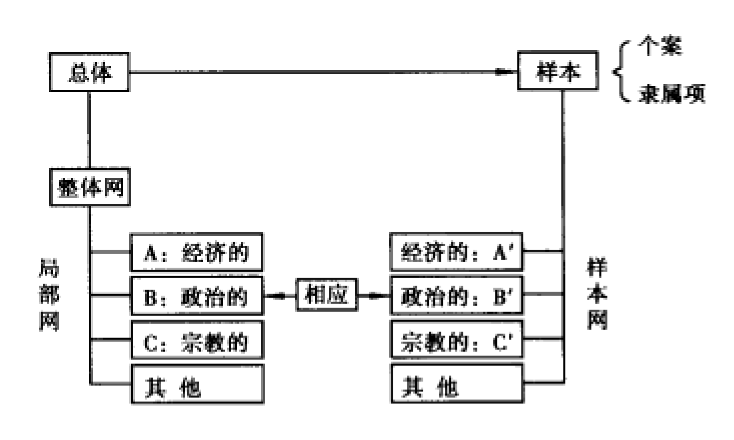


图 3.8 网络和抽样：理想型

抽样的任务应该是显而易见的，也是直接的，涉及的无非是一些在调查研究中的抽样原则：从一个待研究的总体中抽取出有代表性的个案样本，研究它们之间的关系，构造一个样本网络，该网络要与作为一个整体的

总体中出现的局部系统保持相应的关系。但是,实际的问题并非如此简单。抽样的一般原理的基础是基于大量观测基础上的概率论,不仅如此,在判断抽样资料的信度方面,也已经有了一些已得到充分检验的数学规则。但是,在判断来源于一个样本的关系数据的性质方面却不存在这样的规则;我们有充分的理由假设抽样可能导致不可信的数据。尽管我们可能抽出有 1 000 个个案的样本来分析,也可能找到一种用来处理 1 000 × 1 000的邻接矩阵的计算机程序,我们仍然不能保证这个抽样网络的结构会与其对应的局部网结构之间有什么关系。一个有代表性的能动者样本本身并不给出一个有用的关系样本(Alba,1982:44)。

初看起来,这不像是一个真正的问题。例如,能动者之间关系的总体分布以及他们的密度(density)[18]似乎是一个可根据样本数据轻松进行估计的问题:样本提供了每个个案的网络属性数据,这些数据可用来计算全体网络参数。例如,要估计一个国家中的朋友关系的密度,可以询问一个随机居民样本,问他们有多少个朋友。如果样本量足够大,那么这些估计值应该是可信的。但是,几乎不可能超出这些基本的参数来测量网络结构的一些定性方面。

其中的原因与从一个行动者样本中获得的关系数据的稀疏性有关。即使存在一个完美的回答率,并且全体回答者完全回答了所有的问题,回答者提名的许多接触者本人仍将不是样本的成员。这意味着,样本各个成员之间的关系数量将是他们之间的全部关系的一个很小的子集合,并且我们没有理由相信在样本的各个行动者之间确定下来的这些关系本身将是一个针对同样这些行动者之间的**全部**关系的一个随机样本。对于一个非常大的总体,比如对于一项全国范围的研究来说,一个随机样本中的**任何**一个成员几乎不可能与该样本中其他成员之间存在任何类型的社会关系。在一个总量为几百万的总体中随机抽取出的两个人之间存在关系的概率是极小的,以至可以忽略不计。因此,一个研究者很难仅仅根据一个随机样本就对全国性总体的关系结构进行指点。伯特(Burt,1983a)曾经进行过一种粗略的估计,即通过抽样而丧失的关系数据量等于(100 - k)个百分点,其中 k 是样本量占总体的百分比。这样看来,他认为,10% 的样本量将丧失 90% 的关系数据。即使占总体的 50% 的样本也将丧失一半的数据。这种数据的丧失使得在常规的抽样研究中来确定派系、聚类以及其他大量的结构性质等几乎成为不可能的事情。

抽样数据也可能在获得关于所研究的特定个体的关系属性测度方面导致一些困难的出现,如果在调查中存在一定数量无回答数据的时候更是如此。例如,在试图估计一个网络中行动者在社会计量学意义上的知名度

的时候，假设在该网络中知名度非常高的行动者极少，并且大多数人的知名度都不高[19]。由于知名度高者数量很少，一个样本不足以充分包含知名度很高的行动者，因而不能推广到网络中知名度的总体模式。这类似如下述情况，即通过全国范围内的随机抽样调查，对少量的精英或者统治阶级进行研究也出现类似的问题。除非样本量确实很大，否则知名度高者出现的数目不会足够大，而一个非常大的样本将大大抵消了抽样的意义。当然，围绕着该问题的一个方案是利用分层抽样，从而保证知名度高的人被选中的概率也较高。然而，这里明显存在的一个困难是，只有当研究者已经知道总体中知名度的分布的时候，这种抽样策略才能得到执行。

到目前为止，针对这些抽样问题，似乎存在三种不同的回应。第一，放弃那种试图测量社会网络的整体性质的努力，把关注点限制在个人的、个体中心网络上。这种研究策略涉及考察人们无限制地选择出来的人，包括那些在样本中不包含的那些人，并且计算出他们的接触者的密度和其他个体中心网络的特征。由于人们不试图推广诸如总体网络的密度或"紧密联系"结构，因此抽样带来的问题就比其他任何类型的社会研究带来的问题更少。这也是韦尔曼（Wellman，1979），费希尔（Fischer，1982）和威尔莫特（Willmott，1986，1987）在关于友谊和社区的研究中所采用的策略。

第二种反应是利用滚雪球方式。弗兰克（Frank，1978a，1979）认为，研究者应该勾勒出一个初始的案例样本，然后收集该样本成员之间的全部接触者，而不用考虑这些接触者是否为初始样本的成员。然后把这些接触者加入到样本之中，并且利用同样的方式发现这些接触者的接触者。多次持续这种扩展过程，我们将发现越来越多与初始样本成员有间接接触的人。研究者必须决定这种滚雪球持续到何时为止。通常情况下，要持续到加入样本中的额外成员的数量迅速减少的一点为止，因为提到的名字已经在第二次或第三次滚雪球中被提到了。弗兰克指出，这种滚雪球方法可允许人们合理地估计诸如接触者的分布、二方组和三方组的数量等事项。当然，一个滚雪球样本不是随机的样本：实际上，所发现的结构"内嵌于"滚雪球抽样方法本身。但是，为了避免在一个随机样本中可能发现的关系稀疏性，恰恰需要进行这种滚雪球抽样。滚雪球抽样方法有一个假设，即构成抽样网络的网络关联部分代表着该网络的全部其他部分。那么，研究者必须对总体及其关系有所了解，以便对其代表性进行评价。但是，滚雪球抽样方法至少使得试图估计结构的哪些特征将成为这种抽样方法本身的产物成为可能，因而可以在分析中加以控制[20]。

针对抽样问题的第三种反应来自于伯特（Burt，1983a），他提议转而关注对社会网络的一些定性特征。伯特特别关注如何区分出"位置"或者结

构定位，如角色等。在一个网络中，如果假定处于类似结构位置的那些行动者将拥有各种共同的社会属性的话，那么就可能利用有关拥有特定属性的行动者之间的一些类型关系的调查数据来估计在网络中可能存在哪些结构位置。从每个回答者那里必然可以获得有关他们的社会属性方面的信息，也可以获得与他们接触之人（包括样本之外的人）的属性信息。然后，即可把各个行动者归类，使之分组到各个行动者集合之中，每个集合的行动者应该具有一些共同的属性，这些集合可安置到一个集合-集合方阵之中，从而表明各种不同类别的成员之间的关系接触频次。例如，我们可能发现，70%的男性白人拥有男性黑人朋友，而只有20%的女性白人拥有男性黑人朋友。伯特认为，如果研究者已经对一个总体中的全部男人和女人进行了完全列举的话，那么利用这些测度就可以估计各种期望出现的社会角色之间的多值关系。

利用抽样方法来研究大尺度的社会系统，其收效甚微。尽管目前似乎不太可能在样本中发现诸如派系、聚类等性质，我们仍然可以期待，随着网络抽样技术的进一步发展，这种发现会成为可能的（Alba，1982；46；Frank，1988）。

注　释

1 即使它在形状上是“正方形的”，有相同数量的行和列，但它在逻辑上却是长方形的。

2 当然，3-模或一般地讲 *n*-模数据也是存在的，这取决于存在的独立点集的数量。然而，现在还没有什么经过检验的方法能处理这些更复杂格式的数据。

3 实际上，发生阵是一个1-列向量。

4 对于这一争论来说，直接在发生阵上进行操作的计算程序是个例外。有些计算程序分析矩阵的行，有些分析列，其他程序则同时分析矩阵的行和列。在这些情况下，有必要确保将适当的能动者集合作为分析的目标。如果某个特定的程序只分析列（通常确实如此），则需要对矩阵进行转置处理，以便分析在初始阵中被设计为“案例”的能动者。关于该命题（即在案例和隶属项之间的区分是任意的）的某些其他例外情况将在后续章节中得到说明。

5 这里提到的以及贯穿本书的“二进制”数据，都涉及数位（*digits*）的使用，用它来指代一个关系的有无。这里不涉及任何把二进制的值作为关系强度的意图。因此，一个值为3的关系（在2个案例之间有3个共同的隶属关系）用二进制格式来表示仅仅是简单的“1”（关系是存在的），而不是二进制的数字11。通常情况下，这不会引起任何混淆，所以我直接按照通常的惯例，只是谈及二值数据和多值数据之间的区别。

6 如果已经把复杂数据化简为简单数据，就可以返回到初始数据，并把它们转变为一种不同形式的简单数据。然而，如果不参照取值或方向的信息，就不可能把初始无向数据转换为诸如多值或有向数据这样的形式，而这些信息在初始的无向数据中是没有进行编码的。

7 大多数字处理程序都以一种专有的格式保存文件，也将有必要以文本或 ASCII 码的格式保存文件。这样的一个文件包含的只是与字母和数字相对应的编码。参见(Roistacher,1979)。

8 被用来指代数据表中各个单元的数字和字母无需被作为数据的标签使用，尤其是当数据被单独标记和注解的时候更是如此。由于表中的数字和字母简单易懂，所以可以无异议地指代数据表中的任何单元。不幸的是，数据表的一些惯例与矩阵的相反：在指代数据表的单元时候，列要先于行。这就是生活！

9 有些任务可以用诸如 ACCESS 这样的数据库程序来完成，虽然这些程序不特别适合于社会网络分析的目的。重要的是要注意到，关系数据库并不像它表现出的那样是一个关于关系数据的数据库。“关系的”这个词语是在两种不同的情况下使用的。布伦特(Brent,1985)对数据库建构的一些原理给出了很好的论述。

10 STRUCTURE 和 UCINET 要求有描述数据格式的标题行。这些标题行在导出前填加在数据的上方，以便给出数据的行数、列数以及数据类型等信息。

11 在图论中，“线(line)”这个词的意义稍有歧义。例如，公司 1 和公司 2 之间的关系可以被描述成是由三条线组成的(每条线代表一个董事)，或表达为一条其值 3 的线。第二种使用方法适用于多数情况。

12 在无向网络中，“头”和“尾”的区别具有任意性。在有向图中，要从头到尾描述各条线：线路的“头”是社群图中各个点的箭头。图 3.4 显示出与各个行元素对应的尾，“头”与列元素相对应。

13 这些术语在对精英的讨论中得到广泛使用(Scott,1999)。最近，劳曼等学者(Laumann et al., 1983)指出了它们与抽样问题普遍存在着关系。

14 参见斯托克曼等学者(Stokman et al., 1985)在此基础上进行的系统的比较。

15 当进行位置分析并且所有高于某个临界值的案例都选出来进行研究的时候，严格意义上说我们并没有在处理一个样本。所选取的符合条件的案例应被称作准-列举。它不是一个完全列举，因为那些与“切线”外部的关联被忽略了。至于提出在大的公司网络中在数据选择方面如何避免这一问题的提议，参见伯科威茨等人(Berkowitz et al., 1979)的著述。

16 一些董事并不独立构成抽取出来的一个样本，这个循环性(circularity)使得许多常规的统计分析无法应用，因为这些统计分析假定对案例进行概率抽样。

17 附录中讨论的 PAJEK 程序在处理非常大的数据库方面拥有极大的潜力。

18 下一章将充分讨论密度这个概念，它能够根据一个网络中能动者所持有的平均关系数来计算。

19 在社会计量学研究中，声望问题是网络分析中“中心性”测量的一种形式。这种测度将在第五章讨论。

20 当然，如果希望避免偏差的话，那么在滚雪球抽样中最初对回答者的选择就是很重要的。

点、线和密度 4

在上一章中，我考察了用矩阵形式对关系数据进行处理和管理的一些方式。社会网络的许多基本特征都可以通过对矩阵进行直接操作——如转置、加法和乘法等运算——来分析，这些运算都会产生关于网络结构的信息。然而，矩阵代数对于大多数研究者来说是比较复杂的（可以参见Meek and Bradley，1986）。一些专业的计算机程序已经使得网络分析变得更容易，更直接，尽管我们仍然可用矩阵来整理和存储关系数据。现有的一些程序填补了各种各样的分析方法，任何运用该程序的人都必须理解它们是如何运作的。

社会网络分析程序的一个通用框架是对图论进行数学研究，它为描述各种网络及其特征提供了一种形式语言。图论把矩阵数据转换成一些形式上的概念和定理，从而与社会网络的一些实质性特征直接相关。如果说社群图是代表关系矩阵数据的一种方法的话，那么，图论语言则是另外一种更普遍的代表关系数据的方法。尽管图论不是唯一一种用来对社会网络进行模型化处理的数学理论，但它是许多社会网络分析的最基本观念的人手点。

尽管现在的一些唾手可得的计算机程序努力使数学尽量远离该程序的使用者，然而正是图论的一些概念才扮演着 UCINET 和 GRADAP 这两个软件包的主要程序。矩阵形式的数据可以被这些程序读取，相应的一些图论概念也可以得到探讨，而研究者却不必知道其中涉及的理论或者矩阵代数的具体机制是什么。尽管如此，如果对图论有所理解，则会明显地增加研究者分析的精致性，因为他/她会选择恰当的程序。实际上，GRADAP 的数据结构和管理程序也要求对图论的一些基本观念有所理解。

图论关注一系列要素构成的集合及这些要素之间的关系，要素被称为

“点”,关系叫做“线”[1]。这样,一个用来描述群体成员之间关系的矩阵就可以转换成一个由线、点连成的图。因此,一个社群图就是一个“图”。到目前为止,这种讨论与第2章和第3章讨论的内容非常接近。必须明确的是,这种“图”的观念与在统计学以及其他定量数学分析中所用的变量“图”之间有显著的差别,这一点很重要。这些比较熟悉的图形常常被称为“变量图”,例如,数轴上的频次数据表征的是变量。图论中的图是“网络图”,它表达了各点之间关系的定性模式。确实,在图论中,这些图形本身居于次要地位。如前所述,对于具有多个点和复杂联系模式的点集来说,很难划出一个清楚的、可理解的图形来。通过把图的一些性质用比较抽象的数学形式来表达,才可能无需画图,从而容易对非常大的图形进行操作。

无论如何,在图论中,画图表具有极其重要的展示价值,本书也将使用其他图表。由于小图具有简单可视性,因此,在进一步介绍图论的一些基本概念之前,我将首先介绍画图的一些基本原则。

社群图和图论

画一个图表[2]的目的是用纸上的点来表征“发生矩阵”中的每一行或每一列,它们可能代表有待研究的案例或者隶属项。一旦给出了相应的邻接矩阵,矩阵各个格值中的“1”和“0”就代表关系的存在与否,这可以用点与点之间是否存在线来表示。例如,在图3.5中,各个公司之间的4×4对称邻接矩阵可以用一个包含6条线的4点图来表示,这些线对应矩阵中的非0项。

在一个图中,重要的是关联的**模式**,而不是纸上的点的实际位置。图论专家对两点之间的相对位置、连接两点的线的长短、表达点的字母的大小等不感兴趣。图论确实涉及长度和位置的概念,但是这些概念不对应于我们最熟悉的空间长度和位置的概念。在一个图中,全部线的长度一般都画成相等,不管这是否可能,但这完全是一种纯粹的审美约定,是一件完全出于实践上的方便的事情。实际上,如果希望把图画得更精确一些的话,那么这种约定并非总能得以维持。因此,不存在一种直接的画图方式。例如,在图4.1中的各个图都同样有效地表达了同一个图——它们表达的图论信息是完全相同的。

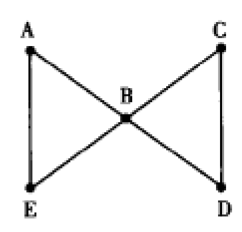
因此,一些图论概念被用来描述各个点之间的关系模式。最简单的图论概念指的是用来构造一个图的各个点和线的诸多性质,而这些概念也恰恰是比较复杂的结构观念得以建构的基础。本章将对这些基本概念进行

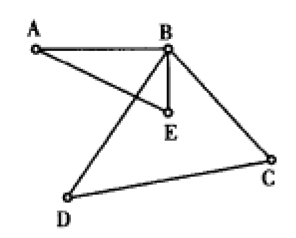
述评，并展示它们如何既针对网络的个体中心性质，也针对整体性质进行概要式考察。下一章将探讨一些更复杂的概念。

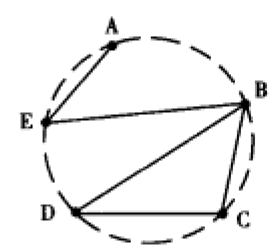
(1)邻接阵

	A	B	C	D	E	行总和
A	-	1	0	0	1	2
B	1	-	1	1	1	4
C	0	1	-	1	0	2
D	0	1	1	-	0	2
E	1	1	0	0	-	2
列总和	2	4	2	2	2	

(2)各种表达同样信息的图



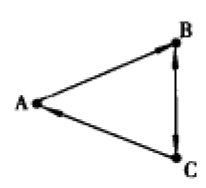




(注:第8章将讨论有关圆形图的约定)

图 4.1 图的各种画法

首先，有必要先看一看在建构图的时候所使用的线的类型。线可以对应图 3.6 区分出来的任何一种关系类型：无向线；有向线；多值线或者有向多值线。图 4.1 中的全部图都由无向线组成。这些图都来源于一个对称数据矩阵，该矩阵仅仅表明了一种重要关系的存在与否。如果一些关系是从一个人指向另外一个人的，那么就用"方向图(directed graph)"来表示关系，有时简称为"有向图(digraph)"。一个有向图要用带箭头的线画出来，箭头的方向代表关系的方向。图 4.2 即为一个简单的有向图。



	A	B	C	行总和
A	-	1	0	1
B	0	-	1	1
C	1	1	-	2
列总和	1	2	1	

图 4.2 有向图及其矩阵

另一方面，如果关系的"密度"很重要并且可以用一定数值来表达的话，那么研究者可以据此建构一个多值图（valued graph），把一定的数值赋予每条线上。我已经指出，一个有向图对应的矩阵通常是不对称的，因为关系通常不是对称的。一个有向图矩阵可能对称，也可能不对称，但是它包含的将是一些数值而不是简单的二值项[3]。图 3.5 表达的就是一个多值图例子。一种最常用、最简单的测量关系密度的测度是一条线的"**多重度**（**multiplicity**）"。它无非是构成关系的独立接触者的数目。例如，如果两个公司共享的董事有 2 个，那么这二个公司之间的关系就用一条多重度为 2 的线来表达。如果两个公司共享 3 个董事，那么连锁董事职位就可以看成是一条多重度为 3 的线。当然，一个图中的各个值能与其他任何恰当的密度测度如关系的频次有关联。

图论的一些基本观念最容易在简单的非多值图中得到理解。人们利用很多直截了当的，明显的词汇来指称图论术语，如果细致地界定这些术语则显得有学究气。但是这些界定是很重要的，因为一些明显简单的词汇是以一些从高度专业化的技术的方式来加以使用的。如果要理解图论的有力之处，那么必须澄清这些词汇的含义，这一点很关键。

一般称由一条线连着的点是相互"邻接的（adjacent）"。邻接是对由两个点代表的两个行动者之间直接相关这个事实的图论表达。与某个特定点相邻的那些点称为该点的"邻域（neighborhood）"，邻域中的总点数称为"度数（degree）"（严格地说应该是"关联度"，degree of connection）。这样，一个点的度数就是对其"邻域"规模大小的一种数值测度。在一个邻接矩阵中，一个点的度数用该点所对应的行或者列的各项中的非 0 值总数来表示。如果数据为二值的，如图 4.1 所示，那么一个点的度数就是该点所在行的总和或者列总和[4]。由于每个图中的每条线都连着两个点——它附属于两个点——所有点的度数总和因而一定是图中线的总数的 2 倍。这是因为，在计算各个独立点的度数的时候，每条线都被计算了两次。这一点可以通过考察图 4.1 而得到确认。在该图中，点 B 的度数是 4，所有其他点的度数都是 2。因此，度数总和为 12，它等于线数（6）的 2 倍。

各个点可以通过一条线直接相连，也可以通过一系列线间接相连。在一个图中的这一系列线就叫做一条"线路（walk）"。如果线路中的每个点和每条线都各不相同，则称该线路为"途径（path）"。途径这个概念是继"点"和"线"之后的另一个最基本的图论术语。一个途径的**长度**（length）用构成该途径的线的条数来测量。例如，在图 4.1 中，点 A 和 D 不直接相连，而是通过途径 ABD 关联在一起，它的长度是 2。图论中的另一个非常重要的概念是"**距离**（distance）"，但是，"距离"和"途径"的概念都不同于

它们在日常生活中的含义。笔者已经指出,一条途径的长度仅仅指它所包含的线数,即从一个点走向另外一个点(所必须经过的"步骤"。而两点之间的"距离"指的是连接这两点的最短途径(即捷径,geodesic))的长度。

请看图4.3。在该图中,AD是长度为1的一条途径(它是一条线),而ABCD是一条长度为3的途径。ABCDA这个线路则不是一个途径,因为它通过点A两次。可以看出,有三条不同的途径把点A和D连在一起:长度为1的AD,长度为2的ACD,长度为3的ABCD[5]。而A和D之间的距离是二者之间的最短途径的长度,也就是1。另外,B和D之间的距离是2。许多比较复杂的图论测度仅考虑到捷径,而其他测度将考虑到图中的全部线。

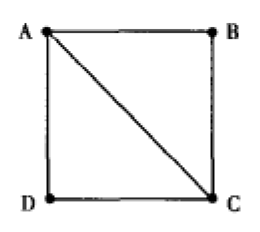


图4.3 线和途径

同样,也可用这些概念来分析有向图,尽管在分析的时候需要作些调整。在一个有向图中,线是指向或者来自不同点的,所以必须考察它的方向。在简单的无向关系数据中,关系是对称的,而有向图中的关系一般不对称。例如,A选择B为友并不意味着B也选择A为友。因此,有向图中某点的"度数"包括两个不同方面,分别称为"点入度(in-degree)"和"点出度(**out-degree**)"。它们的定义要考虑到表达社会关系的线的方向。一个点的点入度指的是直接指向该点的点数总和;**点出度**指该点所直接指向的其他点的总数。因此,一个点的点入度体现在有向图矩阵中该点对应的列总和上,点出度用行总和来表示。例如,在图4.2中,点B的列总和为2,因为它"收到"了两条线(来自于A和C)。其对应的社群图则清楚地显示出它的点入度为2。另外,点B的行总和为1,这反映了它仅"指向"一个点(即点C)。

在一个有向图中,途径指的是箭头指向相同的一系列线。例如,在图4.2中,CAB是一个途径,而CBA却不是:箭头方向的改变意味着C不可能通过B"到达"A点[6]。可见,有向图中用来确定"关联"的标准是相当严格的,因为研究者必须考察到线的方向,而不能仅仅看线的有无。例如,在一个有向图中,两点之间的距离只能根据在考虑到方向之后所确定下来的途径来测量。例如,当把能动者看成是网络中资源或者信息"流动"的"源头"或者"尽头"的时候,在分析网络图的时候,就必须认真考虑到这种有方向的信息。尽管如此,有时候线的方向可以被合法地忽略掉。如果仅仅关注一条线的**存在与否**,而其方向相对来说不是一个重要因素,此时就可以放松通常的严格关联标准,并且只要在任何两点之间存在一系列线,而不考虑其方向的话,就说二者之间存在关联。在这种分析中,我们说存在

一条"半途径(semi-path)"(而非途径)。在图4.2中,CBA就是一条半途径。因此,如果把有向数据看成是无方向的,这就意味着通常情况下针对无向数据的所有测度都可加以利用。

个体中心密度和社群中心密度

在图论中,一个得到最广泛应用的,甚至可能得到滥用的概念是"密度(density)"。这个概念描述了一个图中各个点之间关联的紧密程度。一个"完备(complete)"图指的是一个其所有点之间都相互邻接的图:每个点都与所有其他点直接关联。这种完备性即使在小网络中也极其少见,密度这个概念试图对线的总分布进行汇总,以便测量图在多大程度上具有这种完备性。各个点之间越相关,图的密度就越大。

这样看来,密度依赖于另外两个网络结构参数:图的内含度(inclusiveness)和图中各点的度数总和。内含度指的是图中各类关联部分包含的总点数。换句话说,一个图的内含度等于其总点数减去孤立点的个数。在比较不同的图的时候,最常用的内含度测度表示为关联的点数与总点数之比。这样看来,对于一个拥有5个孤立点的20点图来说,其内含度就为0.75。一个孤立点不附属于任何线,因此对图的密度没有贡献。所以,图的内含度越高,其密度就越大。然而,相互关联的一些点的关联度可能各不相同。有的点与较多其他点相连,有的点则与较少的点相连。一个图中各点的度数越大,该图的密度就越大。为了测量密度,有必要用一个公式把上述两个参数合在一起。这就需要比较一下一个图中实际拥有的线数和一个完整图中应该具有的线数。

图中实际包含的线数是对其内含度和各点度数的一个直接反映。这在简单图中可以直接计算出来,而在复杂图中则必须通过邻接矩阵来计算。任何图中包含的线数都等于各点度数总和的一半。在图4.1中,如我已经指出的那样,行总和或者列总和的一半恰好是6。在该图中可能出现的最多线数可以很容易地根据它所包含的点数计算出来。一个点可能与其他任何点(除了自身之外)相连,因此,一个拥有n个点的无向图中最多可能拥有$n(n-1)/2$条不同的线。计算出来的$n(n-1)$给出的是图中点对的总数,但是连接这些点的总线数是该数目的一半,因为连接A和B的线与连接B和A的线是一条。这样,有3个点的图最多有3条线相连;4点图最多有6条线相连;5点图最多有10条线相连,依此类推。连线数增长

的速度要比点数增长的速度快得多。实际上,对于计算大型网的各种测度来说,这是一个最大的难题。例如,一个包含 250 个点的图中最多可能包含 31 125 条线。

一个图的**密度**定义为图中实际拥有的连线数与最多可能拥有的线数之比,其表达式为$\frac{l}{n(n-1)/2}$,其中 l 代表图中实际存在的线数[7]。该测度的取值范围为[0,1],一个完备图(complete graph)的密度为 1。图 4.4 展示了不同的图的密度:每个图都包含 4 个点,因而最多可能拥有 6 条线。从图中可以看出密度是如何随着内含度和度数总和的改变而改变的[8]。

相连的点数	4	4	4	3	2	0
内含度	1.0	1.0	1.0	0.7	0.5	0
度数总和	12	8	6	4	2	0
连线数	6	4	3	2	1	0
密　度	1.0	0.7	0.5	0.3	0.1	0

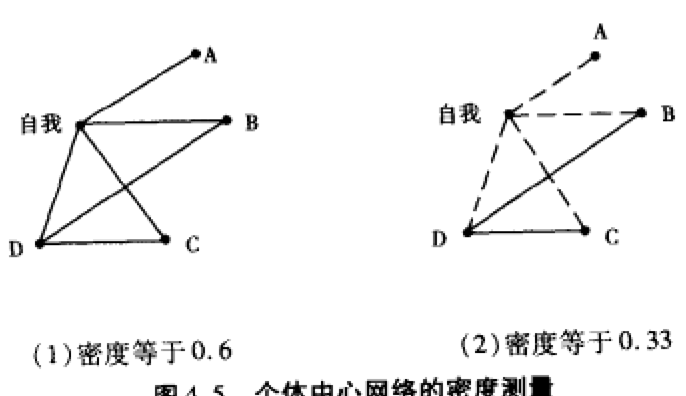
图 4.4　密度比较

有向图密度的计算必然有所不同。有向数据矩阵不对称,因为从 A 指向 B 的有向线不一定包含从 B 指向 A 的线。因此,有向图可能包含的最多连线数恰恰等于它所包含的点的总对数,即 $n(n-1)$,有向图密度的表达式因而是$\frac{l}{n(n-1)}$。

巴恩斯(Barnes,1974)比较了两类社会网络分析。一类是围绕某些特定的参考点而展开的社会网(例如,Mitchell,1969),倡导的是“个体中心(ego-centric)”网研究。从这种视角出发,密度分析关注的是围绕着某些特定行动者的关系的密度。另一方面,巴恩斯也考察了“社会中心(socio-centric)”网研究,关注的是作为一个整体的网络关联模式,这是对社会网络分析的另外一类贡献。从这一视角出发,密度则不再是局部行动者的“个体网”密度,而是整个网络的密度了。巴恩斯认为,社会中心网十分重要,因为一个网络对其成员的约束力不只是通过与该成员直接相连者而起作用。各种间接联系通过关系的构型与那些独立于特定行动者的性质建立了联结,这种联结才应该是关注的核心。

在个体中心网研究中,一个重要问题是:用什么标准对密度进行测量。

在计算个体中心网络密度的时候，通常不考虑核心成员及与该成员有直接关系的接触者，而只关注在这些接触者之间存在的各种联系(links)。图4.5即表示了这样作的结果。社群图(1)表示的是围绕一个“自我”构成的5人网络。该网络表达了该中心者的各种直接联络人以及在这些联络人之间存在的各种关系。该图共有6条线，其密度为0.6。但是这个相对较高的密度主要是因为有4条线把自我与A、B、C、D连在一起。而这些关系几乎要靠界定才能存在，因此应该不予考虑。例如，如果这些数据是通过问卷得到的，即请回答者说出他最要好的4个朋友的名字，那么，这种高密度将是因为“提问”而人为地带来的。回答者与4位被提名者的关系将掩盖这些被提名者之间的关系信息。关于社群图(1)的一个显著事实是，相对来说在自我所接触的各个成员之间的联系很少。在社群图(2)中，与自我直接相关关系用虚线表示，在A、B、C、D之间存在2个关系(用实线表示)，这4人网的密度是0.33。很显然，这才是对主体网密度的一个较有用的测量[9]。



(1)密度等于0.6　　(2)密度等于0.33

图4.5　个体中心网络的密度测量

多值图的密度也是可以测量的，尽管学者们就如何测量这个问题没有达成共识。当然，最简单的作法就是忽略连线上的数值，把多值图简单地看成无向图或者有向图来处理。但是这样作的结果是可能损失大量的信息。例如，有理由认为，多重度高的线对于网络密度的贡献要比多重度低的线的贡献大。这意味着，在一个多值图中，线数应该根据其多重度来加权：一条多重度为3的线就相当于3条线。这样，这种简单的乘法就给出了一个图中实际线数的加权总数。但对于多值图来说，密度公式的分母不容易计算。我们知道，分母指的是一个图可能含有的最多连线数。要确定这个值，首先要知道所研究的网络的最大多重度是多少，这需要有一个假定，而分母值恰恰建立在该假定基础上。如果假定最大的多重度为4，那么连线的加权后的最多个数应该是在与之类似的简单无多值图中线数的4

倍。但是,研究者如何确定某种特定关系的最大多重度的估计值呢? 一个方案是找出在网络中实际存在的最高多重度,并且用它来作为权重(Barnes,1969)。然而,没有任何原因可以说明,为什么实际发现的多重度就应该对应着理论上最大的可能值。事实上,只有当研究者拥有关于所研究的关系之本质的独立信息的时候,多重度的最大值才能被估计出来。例如,在公司董事案例中,平均的董事会规模和董事职位的数目可以作为权数。如果董事会规模平均是5,并且假设任何人拥有的职位都不多于2个,那么一个完备图中的平均多重度就是5。

例如,就图3.5中的公司社群图案例来说,在此基础上测量的连线加权总数将是6的5倍,即30。而该图中实际的连线加权的总数,即把所有线的值加在一起是12,因此,以多重度为基础计算出来的密度就是12/30,即0.4。相比之下,如果不把数据看成是多值的,那么计算出来的密度就是1。然而必须记住的是,以多重度为基础的密度计算要有一个基础,即对一个人可以假定拥有的最多职位数进行论证。例如,如果假定一个人最多可以拥有3个职位,那么该公司社群图的密度将从0.4降低为0.2。对于其他紧密度测度来说,不存在明显的对线进行加权的方式[10]。

因此,多值图的密度测量相对于研究者对数据所作的假定来说是很敏感的。然而,用这种方法计算出来的密度与非多值数据的密度测度完全不可比。所以,研究者不能只简单地应用一种测度,因为这种测度在标准的程序中都有,这一点很重要。研究者必须始终明确在任何程序中包含哪些假定,并且必须把这些假定与计算出来的密度测量值一起汇报出来。如果所说的"多值"指的不是多重度的话,那么对多值数据的处理就更加复杂。

还有一个影响到所有的密度测量的更基本问题必须要考虑到,这就是密度对于图的规模的依赖性问题,它使得不同规模网络的密度难于比较(参见 Friedkin,1981;Niemeijer,1973;Snijders,1981)。我们已经知道,图中包含的线数的不同就造成了密度各异,这一点可以与完备图中的线数作比较。我们有充分的理由相信,任何实际图的最大连线数要少于理论上的最多连线数。如果每个行动者能够维持的关系数有一个上限,那么整个图的总线数就由行动者个数加以限定。对总线数的这种限定意味着,在其他因素保持不变的情况下,大图的密度要比小图的密度小。这一点尤其与能动者可以掌控的时间限制有关联。梅尤和莱文杰(Mayhew and Levinger,1976)认为,人们投入到建立并维持某些关系的时间是有限的。因此分配到某个特定关系的维持的时间就更有限,并且随着接触人数的增加而减少。因此,当回报减少并且代价太大时,行动者就会决定停止建立新的关系,不进行新的时间投入。这样看来,他们保持接触的人数将随着网络规

模的增加而减少。时间上的限制会对接触者的数量因而也对网络的密度产生限制。梅尤和莱文杰利用随机选择模型指出，在实际的网络图中能够发现的最大密度值是0.5[11]。

行动者维持关系的能力也受到关系的特定类型的限制。例如，"爱"的关系一般要比"认识"关系有更多的情感投入，人们当然也意识到自己认识的人比喜爱的人多。这意味着，任何"爱"的关系网要比"认识"关系网有较低的密度。

在第3章中，笔者已指出，密度是可以根据样本资料加以合理估计的网络测度之一。前文已经对密度的测量进行了充分的讨论，现在可以更进一步考察上述观点。对来自大样本资料的网络密度进行测量的最简单最直接的方法是根据样本所包含的个案的平均度数来估计它。对于一个规模充分大的有代表性的样本来说，对平均度的测量就像从样本资料中引申出来的任何总体属性测度一样是可信的，尽管我在前文已经指出了为什么样本资料不能反映关系的全部范围。如果认为估计值是可靠的，就可以应用它来计算网络中的线数。"度数和（degree sum）"——即图中所有点的度数之和——等于估计出来的平均度数乘以图中的案例总数。一旦计算出这个总数，线数就是此值的一半，因而容易计算。由于可能存在的最多线数总可以直接根据总点数计算出来（对于无向图来说，它总等于$\frac{n(n-1)}{2}$，所以图的密度可估计为：$\frac{(n\times \text{平均度数})/2}{n(n-1)/2}$，进而可简化为$\frac{(n\times \text{平均度数})}{n(n-1)}$。

格拉诺维特（Granovetter，1976）更进一步，试图给出一种估计密度的方法，当研究者对初始平均度数的初始估计值的信度尚未把握的时候，可用此方法进行估计。在很多情况下这种估计值具有较高的信度。例如，就公司连锁数据来说，现有的公司信息名录可使研究者获得关于样本公司与总体中的全部公司之间的全部关系信息，当然这要受到各个名录的精确性的限制。在这种情况下，对平均度数的估计就是可信的。另一方面，在研究熟人关系的时候，这种可信性通常并非如此，特别是当总体非常大的时候，可信性更有问题。格拉诺维特的解决方案是放弃单个的大样本，而选择一系列小样本。在每一个子样本中，可以考察各个熟识关系图（随机子图）的密度，并且格拉诺维特指出，随机子图的密度的平均值是对总体网络密度的一个可信的估计。格拉诺维特利用标准统计理论指出，对于一个包含100 000人的总体来说，利用容量在100和200之间的很多样本就可以对总体进行可信的估计。如果样本量为100，需要这样的样本5个；如果容量为

200,2 个这样的样本就足够了[12]。在田野研究中,这些观点又得到了进一步探讨,从而确证了这种一般性策略(Erickson and Nosanchuck,1983;Erickson et al. ,1981)。

这样看来,无论对于无向图还是有向图来说,密度都是一个易于计算的测度,它既可以应用到个体中心网,也可应用于社会中心网络,并且可以从样本数据中被可信地估计出来。毫不奇怪,它已经成为社会网络分析中的最常用的一种测度。尽管如此,我倒希望对它的应用进行限制。对于多值数据来说,它就是一个有争议的测度,它因关系类型的不同、图形规模的不同而不同,因此,不能用它来比较规模差距显著的网络。尽管有这些局限性,对密度的测量仍将在社会网络分析中占据重要地位。如果把它与诸如内含度、网络规模等测度一起汇报出来,那么密度这个概念仍将在社会网络的比较研究中扮演重要角色。

社区结构和密度

我们可以利用一些实际例子来说明密度分析的优势及应用性。作为哈里森·怀特领导的哈佛网络分析小组最早的成员之一,巴里·韦尔曼(Wellman,1979,1982)曾经对社区结构进行了一次大型研究,密度一词在其中起到了关键作用。他首先分析了早已有之的社区研究传统,在该传统中,"社区"研究者一般关注并研究的问题是,与小规模的乡村有关的社区团结是否能够抵抗来自工业化和城市化的各种现代化力量。韦尔曼希望用社会网络分析方法,探讨现代社会的发展是否导致社区的消失和城市反常的出现。一些社区研究的批评者曾经指出,所有类型的社会关系都已经从特殊的领域脱离出来,各种关系越来越具有国家或者国际的性质(参见Bulmer,1985)。韦尔曼试图根据对多伦多的某个城市地区——东约克的研究来考察这个问题,并且与费希尔(Fischer,1977,1982)一样,韦尔曼关注的问题也是"个人社区"是否已经超出当地的邻里界限之外。

东约克是一个内城的郊区,其住房多为私产和公寓单元。在开展研究的1968年之时,它的居民主要有熟练的劳动工人和坐班的白领工人。该田野研究涉及对由随机抽取的845个成年人构成的一个样本进行访谈。在访谈中有一个核心问题是请被访问者说出最常联系的6个人的名字。然后请他们说出这些被提名者之间的关系是否紧密(参见McCallister and Fischer,1978)。对这些问题的回答可用来建构每位被访者的亲密关系个

体中心网。通过询问被每个回答者提名的各个人之间的关系，韦尔曼就能够测量每个个体网的密度。对密度的计算遵循的是前文给出的策略，即要忽略回答者和联络者之间的关系。也就是说，数据是根据“自我”及其6个亲密关系收集到的，但是密度的测量却只根据6个被提名者之间的关系。

韦尔曼发现，许多亲密接触者（大约一半）是回答者的亲属，但是，亲属关系和非亲属关系在各个地区都广泛存在。大多数关系都是东约克人与城里人之间的联系，尽管其中只有很少的关系是东约克人的近在咫尺的关系。在被提名的全部的紧密关系中，有1/4的人住在城外，有的甚至在国外。在给出关于个体社会网的宽泛架构的这些汇总性陈述之后，韦尔曼转而分析这些网络的密度。这些回答者的个体中心网的平均密度[13]为0.33，仅有1/5的网络的密度超过0.5（Wellman，1979：1215）。密度为0.33意味着，在紧密接触者之间可能存在的15个联络中，实际上仅存在5个联络[14]。韦尔曼发现，密度最大的网络主要由亲属构成，这是因为存在如下事实，即只有回答者的亲属才更易于保持相互接触。在不存在亲属义务的网络中，此类接触一般也不易保持。

密　度	此类网络所占百分比/%	成员为亲属的网所占的百分比/%
0.00－0.25	47.1	36.4
0.26－0.50	31.7	56.9
0.51－0.75	7.9	56.9
0.76－1.00	13.3	73.7
	100（n=824）	

图4.6　多个个体网的密度

韦尔曼关于个体网的主要发现可以归结为图4.6。他对上述数据的解释是，人们参与的网络多数是“稀疏连接”的，即密度低。“社区”关系既不是团结性的也不是地方化的（localized）。人们都有一些可以求助的他人，但是个体网的低密度以及相互关系的缺乏表明，这种求助是有限的。尽管如此，这些个体网无论在日常生活中，还是在紧急情况下都是获得帮助和支持的重要来源。“东约克人至少可以获得一个亲友的帮助，但是不能指望得到大多数亲友的帮助”（Wellman，1979：1217）。那些不太给予帮助和支持的亲密者的重要性却表现在交际方面。帮助者常常是亲属，而在社交方面最重要的人物常常是同乡或者同事。

为了深入探讨一些深层次问题，韦尔曼等人在1977到1978年间对最

初的34个回答者进行了深人访谈式的跟踪研究。其目的是为了得到关于前期研究中结构资料的"定性"方面的背景资料。尽管该研究得到的具体结论超出了本章的直接关切点，但是其关注的某些方向可以概要提及。韦尔曼发现，家庭的人际关系网因性别分工和家庭成员参与带薪工作的不同而有所区别。例如，该研究发现，在那些女性挣钱的家庭和女性只作家务的家庭之间有很多差别。一个家庭的社会关系及其获得亲属、朋友、邻里以及同事的帮助主要是通过女人而不是通过男人来实现的。尤其是在女人完全从事家务的家庭里更是如此。而对于那些女性既从事家务又挣钱的家庭来说，他们的社会关系网络的密度很低，因此他们获得的帮助就较少(Wellman，1985)[15]。

韦尔曼的研究所用的关系数据是通过调查分析得来的，但是，同样的观点也可以应用到其他形式的关系数据。例如，史密斯(Smith，1979)利用来自于文献分析的历史数据研究了13世纪英国一个农村的社区模式。史密斯的数据来源于萨福克(Suffolk)①的莱德格里夫(Redgrave)庄园，是一些关于土地占用、财产交易、村民之间的资金争执等方面的数据。他共考察了在1259年到1293年间的575个人之间的13 592种关系。最初，他分析了多种形式的关系及其频次等，从而表明大约2/3的关系是"承诺"(pledging)关系。即在关系到债务偿还以及其他的资金安排的时候，一个人总是给予一定的法律承诺，以便支持他者。

史密斯关注点的问题是，在组建这些关系并且构造这些社区关系的时候，亲属关系和其他的地方(local)关系扮演何种角色。霍曼斯(Homans，1941)曾经对社区团结进行过类似的历史研究，但是他没有用社会网络这个概念。相比之下，史密斯用主体网作为他的主要概念。他根据土地拥有量的多少，把1289年莱德格里夫地区的425个地主分为4类，然后，从每类中等概率地抽取出随机样本，这一过程产生了112个个案，把他们在1283年到1292年这10年间与所有其他人的关系从总数据库中抽取出来。再根据他们的社会基础和地理范围，分析这112个人的个体网(仅考虑距离为1的关系)。每个个人网密度的分布相对于占地来说是曲线关系。对于那些占有的土地不多于4英亩的人来说，密度随着土地的增加而逐渐增长，而拥有4英亩土地以上之人的网络密度却随着土地的增加而减少。因此，有3英亩或者4英亩土地之人的网络密度最大，其密度的中位值介于0.2和0.4之间。他们同时也是最能够整合到多重关系中的人。因此，能

---

① 英国英格兰东部的一个郡——译注。

够最佳地整合到乡村社区中的人恰恰是处于中间层次的地主。考虑到早期讨论的网络规模与密度之间的关系，令人关注的是，史密斯发现，这两个测度之间的关系系数仅仅为0.012。因此，他得出结论：所发现的网络密度的差异不是网络规模的一个纯粹制造物，而是反映了各种人际关系在性质上的真正差异。

在考察到所有网络资料之后，史密斯放弃了下述观念：一个紧密联络的有机社区是围绕亲属和邻里而组织起来的。中世纪乡村的网络结构，至少就莱德格里夫地区而言，要比这种图景松散得多：

> “对于那些经常与近邻互动的人来说，他们也常常与亲属有互动关系，尽管亲属在大多数情况下都远离而居。如下亲属都有亲近的倾向：兄弟姐妹，叔伯大爷，侄子(女)外甥(女)，父母儿女”(Smith，1979：244)。

韦尔曼认识到，他研究的东约克人的个体中心网与通过重叠关系建立起来的多重链条连在一起：他指出，许多网络与流落到“更大的都市网络”中的多个个体网“连接”在一起(Wellman，1979：1227)。但是他没有直接研究东约克人的社会中心网的这些整体性质。而格里科(Grieco，1987)在推广怀特(White，1970)和格拉诺维特(Granovetter，1974)的研究中则提供了此类“连接”研究。格里科关注的是职业机会信息的发送和收取。她的研究表明，某些特定的个体向他们的网络接触者提供的帮助可以改变网络的整体结构。在信息是间接获取的地方，即信息来源于距离不小于2的联络人的地方，存在一种直接关联的趋势，即在信息发出者和信息接受者之间有建立一种新的直接关系的趋势(Grieco，1987：108页及以后)。因此，网络的总密度会提高，其中的某些联络还会通过团结感和义务感而得到进一步巩固和加强。这样看来，密度的初始增长趋势还将保持下去。当网络中的其他人获得了回报他人之助的能力的时候，他们将反过来建立新的直接关系，从而进一步改变网络的密度。这样，个体中心网意义上的多个接触者在个人层次的改变导致了网络密度的持续转变，从而带来该网络的其他社会中心的、总体特征的转变。

注　释

1 有时候，点指的是“顶点”或“节点”，线有时被称作“边”或“弧”。使用这些替换性的词没有什么实质性好处，因此我保持最简单的用法。

2 “图表(graph diagram)”是一个一般性的术语，与诸如前文画出的电线网相反，“社群图”这个术语专用于表示一个社会网络的图。由于本书关注图论的社会学应用，我有时会交替使用这两个术语。

3 有时候，一个多值图被称为“网络”，这非常容易引起误解。应该避免这个术语学问题，因为所有图都应该被视为网络模型。如在第2章中讨论的平衡论一样，有些作者区分出“符号图”，其中的关系具有正向和负向特征。尽管如此，将符号图看成是多值图的一种似乎更合适，这种多值图中的关系有正向和负向这二值之分。从另一方面讲，也可以将它看作是一个由两个简单图组成的复合图，其中一个图由正值组成，另一个图由负值组成。

4 对于一个方阵中的一个特定的点来说，其行总和与列总和相等。如果已知矩阵下半部分的值，那么除非把其余值包含进去，否则不能计算出行总和与列总和。那些接受矩阵下半部分之值作为数据的社会网络分析程序会自动对其进行调节。

5 老师曾经告诉我们，所有的测量都必须有一个属于它们的单位。在图论中，“线”通常就是这样的单位。也就是说，我们可以称两点之间的距离是“三条线”。然而，这个单位通常不是给定的。

6 然而需要注意，点C和A是由一条线直接连接的。

7 如果完成相关公式，实际的线数就等于总度数值的一半。因此密度也可以表达为 $\sum d_i/n(n-1)$，其中 d_i 是点 i 的度数值。

8 内含度（即实际相连的点数与总点数之比）有时可以更有意义地表达为百分比，但这对于小的数字来说并不适合。

9 参见夏基（Sharkey，1989，1990）和蒂姆斯（Timms，1990）近期关于这一问题的讨论。

10 GRADAP程序允许在多样性权重的基础上计算密度，但对于多值图来说它并不提供其他密度测度。

11 在布劳（Blau，1997a，b），赖蒂纳（Rytina，1982）以及赖蒂纳和摩根（Rytina and Morgan，1982）的著述中可以找到关于规模的意义的更一般性讨论。

12 摩根和赖蒂纳（Morgan and Rytina，1977）概述了这种研究中存在的一些问题和局限性，格拉诺维特（Granovetter，1977）已经答复了他们。

13 韦尔曼以百分比的形式给出了一些计算结果，我在本章中将这些结果转换为本章前文定义的基点。

14 韦尔曼假设所有关系都是对称的：如果回答者说好朋友A与好朋友B关系亲密，那么可以假设好朋友B与好朋友A关系也亲密。注意这种分析只处理回答者自己意想的关系，并不是也不必要是密友之间的真实关系。因此，研究工作与最早有关平衡理论的现象学假设是直接相符的。

15 英国的一位学者沃尔伯纳（Werbner，1990）利用亲属网络来研究支持中的互惠性。

5 中心度和中心势

个人或者组织在其社会网络中的中心度如何，这一思想是社会网络分析者最早探讨的内容之一。这个观点最初体现在社会计量学的概念“明星”中，所谓明星指的是那位在其群体中最受欢迎，或者最受人们关注的中心人物。贝弗拉斯（Bavelas，1950）最先对中心度的形式特征进行了研究，并且在他的开创性研究之后，又出现了大量的中心度概念。各种中心度测度广泛出现造成的一个结果是在此领域出现了很大的混乱。而考察一个图中各个点的相对中心度——即所谓“点度中心度（point centrality）”问题，就能够把大多数这种研究结合起来。但是从这种公认的关注点中也分化出一些截然不同的观点。本章将回顾各种点度中心度测度，尤其关注“整体”中心度和“局部”中心度之间的重要区分。说一个点是局部中心点，这指的是该点在其紧邻的环境中与很多点有关联。例如，如果某点有许多直接相关的“邻点”，我们便说该点是局部中心点。另一方面，如果一个点在网络的总体结构上占据战略上的重要地位，我们就说该点是整体中心点。局部中心度指的是局部某点对其邻点而言的相对重要性，而整体中心度指的是该点在总体网络中的战略重要性。

与点中心度测度相关的另一个概念是一个图的“中心势（centralization）”，这两个词常常被混用，因而造成一定的混乱。例如，弗里曼（Freeman，1979）发表的有影响的重要论文既谈到了“点度中心度”，又谈到了“图的中心度（graph centrality）”。如果把“中心度”这一术语严格地限定为点的中心度，而“中心势”特指作为一个整体的图的中心度，那么，曾经出现的混淆程度将大大降低。例如，中心势指的不是点的相对重要性，而是整个图的总体凝聚力或整合度。例如，图可能或多或少地集中在某些点或点集。人们已经给出许多各不相同的程序来测量中心势指数，从而进一

步增加了困扰该领域的混淆程度。在“中心势”这个观念背后隐含着图的结构“中心”,它是中心化的图所围绕的一个点或者点集合。相对来讲,很少有人试图界定一个图的结构中心思想,下面有必要考虑到这一点。

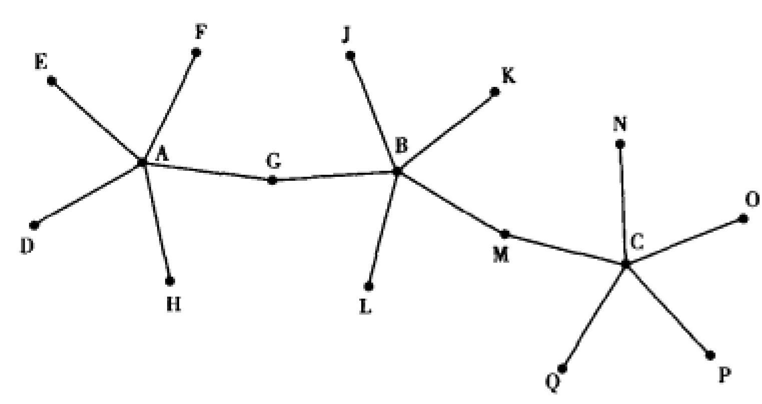
局部中心度和整体中心度

笔者已经指出,点中心度的概念来自于社会计量学的概念“明星”。一个核心点是指那种处在一系列联系的“核心”位置的点,该点与其他点有众多的直接联系。因此,对点中心度的最简单、最直接的测量就是图中各点的度数。所谓度数就是与一点直接相连的其他点的个数。因此,如果某点度数高,则称该点居于中心,在有“关联紧密”或者“处在万物之中”的意义上,该点所对应的行动者就是中心人物。以度数为基础的这种对点中心数的测量对应的是如下直观观念:一个点在其局部环境内是如何被联络的。由于这种测量仅仅根据与该点直接相连的点数,忽略间接相连的点数,因此,所测量出来的度可以称为“**局部中心度**(local centrality)”。对此概念进行系统解释的是聂米南(Nieminen,1974)。以度数为基础的对局部中心度的测量也可以用来计算有向图中点的中心度,尽管在这种情况下每个点都有两种局部中心度测量,一种对应的是点入度,另一种对应点出度。因此,在有向图中有必要对各个点的内中心度(in-centrality)和外中心度(out-centrality)进行区分(Knoke and Burt,1983)。

这种以度数为基础的点中心度测度也可以超越直接联系,达到对各种途径距离的测量。在此情况下,“邻点”的含义就扩大了,从而包含与远距离点之间的关系。这样,对一个点的局部中心度的估计既可以根据直接关系(距离为1),也可以根据距离为2的关系,或者大于2的关系来测量。在利用距离大于2的点来测量一个点的点度中心度的时候,存在的主要问题是,即使在一个密度适中的图中,大多数点都倾向于通过距离相对较少的一些间接点联络在一起。这样看来,如果大多数点与其他点之间的距离都是4的话,那么对各个距离为4的各个局部中心度指数进行比较不会提供更多的信息。距离为1和2的关联才最能够提供信息。显然,选择利用哪个临界点,这完全是研究者的非正式判断的事情,但是距离为1和2提供的信息在大多数研究中都最多。

在对局部中心度进行测量的时候,不涉及整个网络是否有独一无二的“核心”点这样的问题,认识到这一点很重要。例如,在图5.1中,点A,B,C

都可以看成是局部中心点，因为它们的度数都是5，其他点的度数则仅仅为1或者2。即使点A比B和C具有更多的直接关系，它也不是整个网络的中心，因为它在位置上处于由各个点构成的长链的“一侧”，它的中心度完全是“局部的”现象。其度数就是对局部中心度的测量，对一个图中各点的度数的比较表达了各个点在多大程度上与其局部环境中的其他点联络在一起。



	A,C	B	M,G	J,K,L	其他所有点
局部中心度 绝对 5	5	2	1	1	
局部中心度 相对 0.33	0.33	0.13	0.07	0.07	
整体中心度	43	33	37	48	57

图5.1 局部中心度和整体中心度

然而，这种对局部中心度的测量存在着一个重要局限性，即中心度数仅仅在同一个图的成员之间或者在同等规模的图之间进行比较才有意义。除了其他因素之外，一个点的度数依赖于图的规模，因此，当图的规模不同的时候，不同点的局部中心度是不可比较的。如此看来，使用初始度数值可能引起误导。例如，在一个有100个点的图中，度数为25的核心点就不如一个有30个点的图中度数为25的点那样居于核心地位，而这两点都不能轻易地与10点图中度数为6的中心点作比较。为了克服这一问题，弗里曼（Freeman，1979）提出了一种局部中心度的相对测度，它指的是点的实际度数与可能联络的最多度数之比。因此，一个有100个点的图中，度数为25的核心点的相对中心度就是0.25，一个有30个点的图中度数为25的点的相对中心度就是0.86，在一个10点图中度数为6的点的相对中心度[1]为0.66。图5.1也表明，相对中心度也用于比较同一网络中的各个点。还应该明确的是，这种观念也能推广到有向图中。如此看来，相对测度就

是一个测量局部中心度的更加标准化的量。

由初始点度中心测度引起的比较问题与比较不同图形的密度问题有关,这一点已经在上一章进行了讨论。二者都受到图规模的限制。然而,前文已经指出,网络密度水平还依赖于所分析的关系类型。例如,"熟人"关系网的密度要比"喜爱"关系网的密度大。由于密度和点度中心度都要根据度数来计算,因此,对各种点中心度测度都要进行完全同样的思考。例如,如果其他情况保持不变,那么一种"爱"的关系网中的中心度要比熟人关系网的中心度低一些。对相对点度中心度的测量无助于解决此问题。即使利用弗里曼提出来的相对性术语计算出了局部中心度指数,它们还是只能用于同等规模关系网络的比较。

然而,局部中心度仅仅是对点中心度的一类概念化罢了,弗里曼(Freeman,1979,1980)又提出了一种"**整体中心度**(global centrality)"测度,它根据的是各个点之间的"接近性(closeness)"。不管根据什么样的途径距离进行测量,"局部中心度"测度都要根据与某一点直接相联的点数来测量。而弗里曼对"整体中心度"的测量根据的是不同点之间的"**距离**"。前文已指出,两个点是由一条包含不同线的途径(path)连在一起的,途径的长度用组成该途径的线数来测量。在图论中,两点之间的距离用这两点之间最短途径的长度来测量。在地球表面,两点间的最短距离位于连接两点的捷径(geodesic)上。因此,相比之下,在一个图中任何两点之间的最短距离也称为"捷径"。如果一个点与其他许多点的距离都很短,则称该点是整体中心点。在图中,这样的点与图中许多其他点都"接近"。

对接近性的最简单的测量可能是计算"距离和",即图中该点与其他各个点之间的捷径距离之和(Sabidussi,1966)。如果已经计算出来一个无向图中各个点之间的距离矩阵,那么一个点的"距离和"便是该点所在"行和"或者"列和"(二者相等)。一个"距离和"比较低的点与其他许多点都"接近",因此,接近性与距离和是反向的。当然,在一个有向图中,途径必须根据具有相同方向的各条线来测量,因此,根据"行和"与"列和"计算出来"接近性"将有所不同。这样,一个有向图的整体中心度便可以根据所谓的"内接近性(in-closeness)"和"外接近性(out-closeness)"来计算。

图 5.1 比较了整体中心度的"距离和"测度以及各种建立在度数基础上的绝对局部中心度和相对局部中心度测度。由表可见,A,B,C 的局部中心度相同,但是点 B 比 A 和 C 都具有整体中心度。G 和 M 的整体中心度比 B 的小,但是比拥有较高局部中心度的 A 和 C 这两点的整体中心度都高。因此,在"距离和"基础上得到的这些差异也进一步确证了从对图的视觉考察中得到的印象。在对中心度较低的点进行测量的时候,这一点也明

显可见。其他点的度数都是 1,表明具有低局部中心度。然而通过“距离和”测量可以清楚地看出如下事实:J,K,L 的整体中心度要比其他度数为 1 的点的整体中心度高。

弗里曼(Freeman,1979)又增加了一个新的点中心度概念,他称之为中间度(betweenness)。该概念测量的是一个点在多大程度上位于图中其他点“中间”:一个度数相对比较低的点可能起到重要的“中介”作用,因而处于网络的中心。例如在图 5.1 中,点 G 和 M 处于大量“点对”中间。一个点的中间度测量的是该点对应的行动者在多大程度上成为“掮客”或者“中间人”,能在多大程度上控制他人[2]。因此,可以把 G 看成是在以 B 为中心的各个行动者和以 A 为中心的各个行动者之间的中间人,而 M 也在 B 和 C 的各行动者之间起到了同样的中间人作用。

弗里曼对中间度的研究是围绕“局部依赖性(local dependency)”这个概念建立起来的。如果连接一个点同其他一些点的途径经过某点,则称前一点依赖于后一点。伯特(Burt,1992)用“结构洞(structural holes)”概念对此进行了描述。当两个点以距离 2(而不是 1)相连的时候,就说二者之间存在一个结构洞。结构洞的存在使得第三者扮演经纪人或者中间人的角色。例如,在图 5.1 中,点 E 只有依赖于点 A 才能与图中其他点建立联系,点 E 也要依赖于点 G,B,M 和 C,尽管依赖程度有所减少。

在针对点中心度的各种测度中,中间度可能最难计算。点 Y 相对于某一对点 X 和 Z 的“中间度比例(betweenness proportion)”定义为:经过点 Y 并且连接 X 和 Z 这两点的捷径占二者之间总捷径数的比例。它测量的是 Y 在多大程度上位于 X 和 Z“之间”[3]。点 X 对点 Y 的“点对依赖性(pair dependency)”之值就定义为 Y 相对于所有包含 X 的点对的中间性比例之和。“局部依赖矩阵(local dependency matrix)”中包含了这些点对依赖值(pair dependency scores),矩阵各项值表达了每行因素对每列因素的依赖性。一个点的总“中间度”就用该矩阵各列值之和的一半来计算,也就是各列所代表的点对依赖值之和的一半。尽管算起来比较复杂,但是这一测度仍然具有直观意义,并且在 UCINET 和 GRADAP 程序上可以很容易地算出来。

在弗里曼的著述中有针对一整套点中心度进行测量的详细说明:局部中心度(点的度数)、中间度和整体中心度(接近性)。我已经指出,如何通过计算局部中心度(而不是绝对中心度)来推进不同社会网络之间的可比性,弗里曼也在其他形式的中心度测度中指出了这一点。他还提出了相对中间度测度,并且利用比彻姆(Beauchamp,1965)的一个公式提出一种相对接近度测度。然而,所有这些测度都依赖于初始的度数值和距离值,这样,

我们就有必要转向伯纳西茨(Bonacich,1972,1987)的研究,他利用一些权数值提出了另外一类研究。

伯纳西茨指出,对某点的中心度的测量不能脱离所有其他与之相关的点的中心度。与某些中心点相连的点的中心度也会提高,相应地,它也提高与自己相连的其他点的中心度(Bonacich,1972)。因此,在计算中心度的过程中涉及内在的循环。按照伯纳西茨的见解,在一个图中,点 i 的局部中心度 c_i 要根据如下公式来测量:$\sum_j r_{ij}c_j$。这里的 r_{ij} 是连接点 i 和点 j 的线的取值,而 c_j 是点 j 的中心度。也就是说,点 i 的中心度等同于与该点相连的线的取值,并且根据这些点的中心度进行加权[4]。

随后,伯纳西茨(Bonacich,1987)进一步推广了自己的研究。与弗里曼一样,他也给出了一系列局部测度和整体测度。他给出了最一般的中心度测量公式 $c_i = \sum_j r_{ij}(\alpha + \beta c_j)$。在此公式中,中心度权数要根据两个参数 α 和 β 进行修正。α 是一个标准化常数,引入它只是为了保证最终的各个中心度值围绕平均值 1 变动。另一方面,β 更具有实际意义。它有正负之分,允许研究者设置各种不同的途径距离,从而用于计算中心度[5]。当 β 被设定为 0 的时候,就不考虑任何间接关系,中心度的测量因而仅仅是对以度数为基础的局部中心度的测量。β 如果增加,途径的长度也增加,在计算中心度的时候,就逐渐考虑比较远的关系。伯纳西茨声称,以正 β 值为基础的测度与弗里曼对接近度的测量高度相关。

然而,在伯纳西茨的论证中存在的一个重要难题在于,赋予 β 的值完全是研究者自己主观选择的结果。我们很难知道,出于什么样的理论原因使得研究者选择某个 β 值而不是别的值。尽管初始的伯纳西茨测度可能易于直观理解,一般化的模型则难以对大于0的 β 值进行解释。另一方面,β 值可正可负,这确实也提供了分析符号图的途径。伯纳西茨本人建议,负值对应的是一些“零和(zero-sum)”关系,诸如那些在持有资金或者其他金融资源中涉及的关系一样。另一方面,正值对应着一些“非零和(non zero-sum)”关系,如那些在获取信息方面的关系。

笔者已经根据图中最核心点对中心度进行了讨论,但是应该明确的是,根据各个中心度的取值也能区分出最不核心的点。对于那些中心度最小的点来说,不管其中心度是如何测量的,都可以看成是一个图的边缘点。例如,对于图 5.1 中的中心度为 1 的全部点来说确实如此。它们居于局部的边缘地位,与网络是松散关联的。然而,图 5.1 中的各个整体中心度数值表明,点 J,K 和 L 并不像其他度数为 1 的点那样在整体上居于边缘地位。

中心势和图的中心

到目前为止,我的关注点集中在点的中心度问题上。但是,我们还可以考察作为一个整体的图在多大程度上具有一个中心化的结构。密度和中心势(centralization)这两个概念代表的是一个图的总体"紧凑性(compactness)"的不同方面。密度指的是一个图的凝聚力的总体水平;"中心势"描述的则是这种内聚性能够在多大程度上围绕某些特定点组织起来。因此,中心势和密度是两个重要的、彼此相互补充的量度。

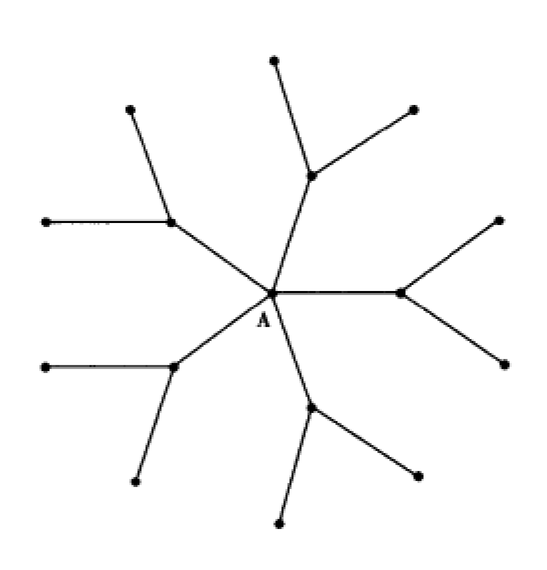


图 5.2 一个高度中心化的图

图 5.2 表示的是一个高度中心化的图的简化模型:整个图以 A 为中心点。如何测量这种层次的中心势?弗里曼(Freeman,1979)已经指出如何把各种点度中心度测度转化为对在不同图中发现的总体层次的中心势水平进行测量。一个图的中心势测度表达了该图围绕某个最核心点的紧密性如何。弗里曼对中心势的测量试图把各种简化的中心势观念的各个方面独立出来。在此基础上,他给出了 3 种类型的图中心势测度,这些测度都根基于他所界定的各种点中心度思想。

无论哪类中心势测度都要用到下面的一般性程序,即找出最核心点的中心度和其他点的中心度之差。这样,**中心势**就是实际的差值总和与最大可能的差值总和之比。弗里曼讨论了 3 类针对这种测度的操作化方法,它们都基于对三种点中心度概念的应用。弗里曼(Freeman,1979)指出,这三种测度的取值都介于 0 和 1 之间,如果一个图的结构是以"星形(star)"或者"辐射形(wheel)"形式组织起来的,那么这三种测度值都为 1。他进一步指出,对于一个完备图来说,这三种测度值都是 0。大多数实际社会网络的中心度都在这两个值之间,并且正是在这些情况下,选择一个还是多个测度对于表明图的特殊结构特征来说就具有重要意义。例如,以度数为基

础的图中心势的测量似乎对各个点的局部重要性很敏感，而一个以中间性为基础的测量则对各个点构成的“链(chaining)”敏感。

评价一个图围绕某个中心点的集中趋势，这是对图的中心趋势进行广义理解的开始。各种中心势测度可以告诉我们一个图是否围绕它的一些最核心点建立起来的，但是没有告诉我们这些中心点是否构成了一个独特的点集，这些点是否成为图中特定的一个聚类(cluster)。例如，在图中各个核心点可能广泛地分布在图的各个部位，在此情况下，对图中心势的测量就可能不会增加我们对图的认识。因此，有必要考察图中是否存在一种可以辨别出来的“结构中心(structural center)”。一个图的**结构中心**是单个点或者一个点集，它像圆心和球心一样，是图结构的枢纽。

这种被称为“核心中心势(nuclear centralization)”的研究在斯托克曼(Stokman)和斯奈德兹(Snijders)[6]的未出版的著作中已经勾勒出来了。他们的研究工作就是把具有最高点中心度值的那些点集界定为图的“中心”。在确定这个点集之后，研究者就可以考察该点集与图中所有其他点之间的关系结构。图5.3表达了二人研究的概要式框架。

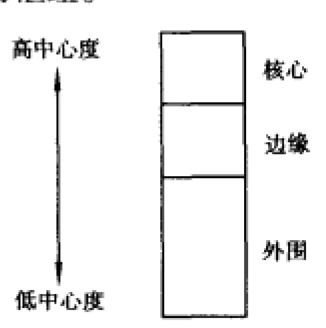


图5.3 图的结构中心

如果图中所有点都按照点中心度——斯托克曼和斯奈德兹用的是局部中心度——的顺序排列起来，那么，具有最高中心度的一系列点就是图的中心。只要在各个中心度值的分布中存在一个“自然的断裂”，就可以画出中心和图中其他点之间的界限来。例如，每一个后续点的中心度的降低可能表明某个点在分布中存在一个跳跃，这就可以看成是中心和“边缘”之间的界限。边缘指的是一系列与中心接近的点聚类，反过来，这些点聚类又与“外围”点区分开来，区分方法也要通过在中心度值分布中的进一步“断裂”。

斯托克曼和斯奈德兹给出的概念仅适用于中心趋势高的图。例如在图5.2这样的以一系列点为中心的图中，如弗里曼给出的一种测度所显示的那样，如果区分出斯托克曼和斯奈德兹界定的点集，那将带来极大的知识增量。当然，在区分中心、边缘和外围的时候会出现不可避免的主观臆断。解决这两个问题(尽管这不是斯托克曼和斯奈德兹关注的问题)的办法是运用某种“派系分析”或者聚类分析，以便确定结构中心的界限：例如，如果最核心的一些点构成了一个界定清晰的、界限明确的“派系”，就有理由认为它们构成了图的核心[7]。但是并非所有图都有这样一个中心点集的等级结构。如果一些中心点并不聚在一起，就不能成为一个中心趋势图的

核心，那么斯托克曼和斯奈德兹的"核心"仅仅由一些局部中心（尽管是分散的）点构成。在这种情况下，用"中心"一词就没有什么意义了。

我们还可以稍稍扩展中心趋势分析，进一步考察图是否可能存在"绝对中心（absolute center）"。一个图的**绝对中心**与圆心或者球心很相似，是图得以构建的那个核心点。作为一个点的**集合**，图的结构中心不满足这个定义。图的绝对中心必须是**单个**点。例如，一个圆心是与圆周上所有点都等距离的唯一点。与之严格对应，一个图的绝对中心与圆心一样，应该与图中其他点的距离相等。这个观点在图中很难操作化，一个自然的想法就是放松该标准，用"最小距离"观念来代替"等距离"标准。这就是说，绝对中心指的是在途径距离的意义上与图中所有其他点最"接近"的一个点。

克利斯托弗兹（Christofides，1975：第 5 章）建议利用距离矩阵对图的绝对点进行概念化处理和计算。与弗里曼提出的方法一样，在他的论证中，第一步也要测量"接近性"。在构造距离矩阵（该矩阵给出了每对点之间的最短途径距离）之后，他定义了点的"**离心度**（eccentricity）"或者"分离性"，并把该值作为矩阵中该点所在列（或者行）的最大值[8]。因此，一个点的离心度指的是与该点相连的最长捷径的长度。克利斯托弗兹对绝对中心度观念的最初估计就是把具有最低离心度的点称为"绝对点"。在图5.4中，社群图（1）的 B 的离心度为 1，所有其他点的离心度为 2。所以点 B 为此图的绝对中心点[9]。然而在其他图中可能不存在具有最小离心度的单个点。在其他图中可能存在许多具有相同的低离心度的点，在这些情况下，就要进行确定绝对中心点的第二步了。

在确定图的绝对中心点的时候，第二步是找出该图中可能具有最小离心度的一个**想象之点**（imaginary point）。也就是说，这里的关键命题是，尽管图的绝对中心点可能在图的某个构成途径上找到，但是这个位置可能不对应图中某个**实际**点。任何图都有一个绝对中心点，但是在某些图中，该点可能是想象之点，实际上不存在。

上述观点不会如初看起来那样令人奇怪。在图 5.4 的（2）中，所有点的离心度都是 2，因此，它们具有相等的"中心性"。然而可能找出一个想象之点 Z，它像在社群图（3）中的情况一样，处在点 A 和 B 之间。点"Z"与 A 和 B 的距离都是 0.5，与点 C，D，E 的距离都是 1.5。想象之点 Z 比任何实际的点都处于"中心"地位，因为其离心度是 1.5。但是我们还是不能找出该图的一个绝对中心点。实际上，想象之点 Z 可以被设置在图中任何一条线的中间点上，其效果是一样的。对于想象之点来说，任何其他位置都不能增加它的最低离心度。因此，对于该图来说，我们最好可以说绝对中心点可能存在于 6 个地点，但是它们都不对应一个实际存在的点。这样看

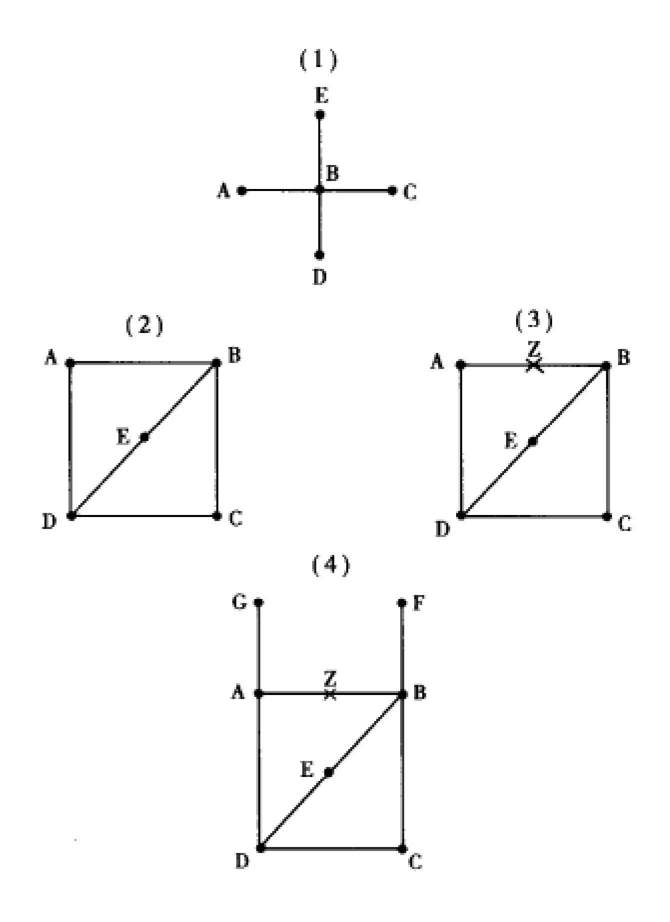


图 5.4 图的绝对中心

来,第二步(即寻找想象点作为绝对中心点)会降低不存在绝对中心点的图的数量,但是它不保证在所有图中都可以找出唯一的绝对中心点[10]。

这样看来,有的图可能存在唯一的绝对中心点,有的图可能存在多个绝对核心点。克利斯托弗兹给出了一种算法,利用迭代法来区分一个图是否包含一个中点,或者实际存在一个绝对核心点[11]。例如,在图 5.4 的社群图(4)中就存在一个绝对核心点,与其他"想象的中点"的离散度为 2.5,离心度为 2 的点 A 和 B、离心度为 3 的 C,D,E,F,G 比较而言,其中"点"Z 的离心度为 1.5。

关于绝对密度的题外话[12]

如前所述，现有的各种对网络密度的测量都与网络规模有关。对于规模差距很大的不同图的比较来说，密度是很难利用的一种测度。规模与密度相关，这就提出了一个问题：是否可能找到绝对密度测度，使之可用于比较研究。笔者在这里不能给出一个完整的答案，但是，图的绝对中心这种思想确实产生了一种可能，即有可能为了测量绝对密度而沿着同样的思路产生出其他一些概念。例如，在物理学中，测量固体物质的密度就需要诸如"半径"、"直径"、"圆周"等概念，所有这些概念都依赖于绝对中心这种思想。

圆形物体或者球形物体的半径指的是中心到边界的距离，在此基础上建立其最远的可达点。把它翻译为图论术语就是，一个图的绝对中心点的离心度可以看成是图的"半径(radius)"。图的"直径"(下一章将详细说明)定义为图中任何点对之间的最大距离。例如，图5.4中(4)的半径为1.5，直径为3。在这种情况下，直径是半径的二倍。这与传统的几何学中圆或者球的情况是一致的。但是，对于所有图来说，情况并非总是如此。

在几何学中，物体的面积和体积之间存在明确的关系，这些关系可以推广到三维以上的物体。一个圆的面积是 πr^2，球的体积是 $4\pi r^3/3$，π 是圆周率。因此，一个圆面积的一般表达式为 cr^2/d，球体积的一般公式是 $4cr^3/3d$，其中 c 是圆周长，r 是半径，d 是直径。把该公式应用到图5.4中的简单社群图(4)中可以得到，它的体积[13]为 $4c(1.5)^3/9$ 或者 $1.5c$。但是在此公式中 c 取什么值？如果把图的直径看成是两个最远点之间的捷径长度(即最长的捷径之长度)，那么周长自然应被看成是图的最长途径(path)。在社群图(4)中，该长度就是连接点G和F的长度为5的途径。此时，该例子中图的"体积"为7.5。

因此，相对简单的几何学促使我们转向在三个维度上对图的绝对中心度进行测量。在物理学中，密度被定义为单位体积中物质的质量，因此需要对图的"质量"进行测量。在物理学中，质量就是一个物体所包含的物质的数量，因此，最直接的图论意义上的"质量"概念就是图所包含的线数。在社群图(4)中有8条线，所以其绝对密度就是8/7.5，或者1.06。

推而广之，我们可以给出测量一个图的**绝对密度**的公式：$l/(4cr^3/3d)$，这里的 l 是线数。与上一章讨论的相对密度测度不同，这个公式给出的是

绝对值,因此可以与任何规模的图进行比较。但是也必须加入一个重要的保留意见,即绝对密度的值依赖于测量时涉及的维数。上述测量是在三维图上进行的。这个概念可以推广到更高维度上,这要利用"超体积(hyper-volumes)"公式,但是这种研究要求在如何决定一个图的维数上达成共识。笔者将在参考弗里曼(Freeman,1983)[14]的研究基础上,在第8章对该问题作进一步分析。

公司网络中的银行中心性

对各个公司之间的连锁董事的研究虽然不是新课题,但是1970年代以前的绝大多数此类研究很少正式应用社会网络分析方法。尽管对密度测量和聚类(cluster)分析的应用是有限的,大多数此类研究坚持严格的定量思路,仅仅计算各个公司的董事数以及公司之间的连锁董事数。莱文(Levine,1972)发表的影响深远的论文标志着这种研究的转向,与此同时,莫肯及其同事在荷兰也进行着一项开创性研究,他们全面系统地应用图论来探讨连锁董事(Helmers et al. ,1975)。然而,主要转折点发生在1975年,当时迈克尔·施瓦兹(Michael Schwartz)和学生提交了一篇会议论文,该文应用中心度概念于公司网络之中(Bearden et al. ,1975)。这篇长文虽未正式发表,却被广泛传阅,其影响是深远的。施瓦兹小组的工作以及在他们的激励下出现的其他研究,为点中心度观念具有的概念力量提供了令人信服的例示。

迈克尔·施瓦兹和彼得·马里奥里斯(Peter Mariolis)在1970年代初就着手建立美国大公司的数据库,二人的工作为许多后续研究提供了大量的数据(例如参见 Mariolis,1975;Sonquist and Koenig,1975)。他们后来逐渐扩大数据库,追加了1962年美国工业500强,金融和商业250强数据,以及从1963年到1973年间每年所有新加入到这"750强"公司的数据。最后的数据库包含了在1962—1973年间美国1 131家最大公司的所有董事名单;总数为13 574。无论从哪个角度上说,该数据库都是一个大社会网络。如此看来,需要根据具体年限选择出具有实质性意义的子数据库。一个这样的子数据库就是1969年的797个最大企业群体的数据,马里奥里斯(Mariolis,1975)的研究使用的就是这个数据。

施瓦兹及其同事的开创性论文(Bearden et al. ,1975)使用了施瓦兹-马里奥里斯(Schwartz-Mariolis)数据库,并且利用格拉诺维特(Granovetter,

1973）对强、弱关系之间的概念区分的思想对这些数据进行了分析。他们的论断基础是，各个公司专职职员之间的连锁（interlock）可以看成是公司网络中的“强”关系，而那些兼职的、非行政官员之间的联系则是“弱”关系。这一理论论断的基础是，专职行政领导之间的关系最可能是董事会层次的联系，对相关公司来说具有战略意义。因此，他们倾向于持有多个公司的股票，从而在公司之间建立了贸易关系[15]。另一方面，非执行董事之间的连锁牵涉的时间投入较少，对相关公司的战略意义较少。

人们应用伯纳西茨（Bonacich，1972）的测度考察了各个大公司的中心度。如前所述，某个特定点的中心度需要根据该点的度数，与该点相连的每条线的值，以及与该点相连的其他点的中心度的组合来测量。这是一个“循环”测度，因而需要大量运算。例如，一个包含 750 个公司的网络需要解 750 个联立方程组（simultaneous equations）。在比尔登（Bearden）等人的分析中，第一步是决定用什么测度测量连接各个公司之间线的值。对于弱的、无方向线来说，比尔登等人认为，每条线的值应该仅仅是独立连锁关系的数量，并且用两个董事会的规模进行加权。该权数依赖于一个假设，即拥有大量的连锁关系对于董事会规模大的企业来说，其意义不如对于董事会规模小的企业的意义大。其计算公式是 $b_{ij}/\sqrt{d_i d_j}$，其中的 b_{ij} 是在公司 i 和公司 j 之间的连锁数，d_i 和 d_j 分别是这两个公司的董事会规模。这个公式允许在图中所有“弱”关系的基础上对伯纳西茨的中心度测量进行计算。

如果根据强关系对中心度进行测量，则需要比较复杂的公式。在此情况下，对每条线取值的测量则要考虑到图中每条线的“方向”。对于作为“发送”线［在 GRADAP 程序中称之为“尾（tails）”］的公司来说，线的取值根据董事“发出”的关系来计算，并且根据作为“接受者”的公司的董事会规模来加权。此程序试图根据连锁关系相对于作为接收者的董事会的重要性来对线进行加权处理。反之，对于那些连锁关系的“接受者”公司（“头”）来说，所收到的连锁关系数要根据作为发送者的董事会规模来加权[16]。比尔登等人认为，最后对中心度的测量要引入更进一步的权重。不直接采用“尾”和“头”的初始加权值，而是采用发送者值的 90%，接受者值的 10%。在这种加权值背后的推理在于一种理论判断，即在公司联络领域中，“‘给予’比‘接受’更重要”：输送一位董事的企业常常标志着该企业比接受董事的企业更有权力。因此，针对各个中心度值引入的人为调整就是为了在最终结果中体现出这个理论判断。然而，需要注意的是，中心度不总是表达权力的一个符号。在某些情况下，最主要的、最可见的行动者可能恰恰处在最弱者中间（Mizruchi，1994：331-2）。

在该研究中，用伯纳西茨的中心度测量方法计算出来的公司中心度值

与公司的度数高度相关，相关系数为0.91。然而，比尔登等人认为，最好选用比较复杂的伯纳西茨测度来测量，因为它有可能突出那些度数低、但是却与高度核心企业相关的公司。他们认为，这种位置对于决定这些公司在经济中的结构显著性来说非常重要。

施瓦兹和同事对中心度作进一步研究，他们称之为"顶点分析(peak analysis)"。后来，米祖齐(Mizruchi, 1982)进一步把"顶点分析"看成是对美国的公司网络在20世纪的发展进行解释的基础。一般认为，如果一个点比任何与该点相连的其他点更处于中心地位，则称前者为"顶点(peak)"。明兹和施瓦兹(Mintz and Schwartz, 1985)推进了这一思想，提出了把两个或多个"顶点"连在一起的"桥接点(bridge)"这个概念(参见图5.5)。他们进一步把"聚类(cluster)"看成是由所有与"顶点"直接关联的点构成的，但是却要去掉那些与另外一个顶点距离为1的点。这样，顶点就处于这些聚类的核心[17]。

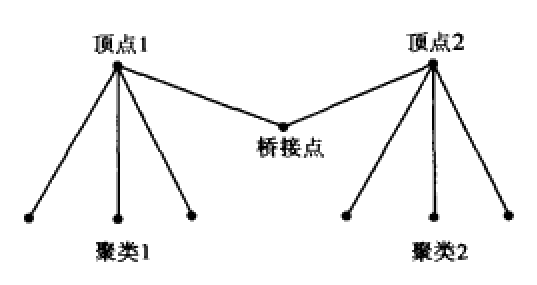


图5.5 顶点和桥接点

通过运用这些技术来测量点的中心度，得到的结论被广泛接受，这些结论也表明了公司网络的一些最基本的特点。总之，比尔登等人认为，美国的公司关系网总体上表现为具有"银行中心度"模式：在网络中，无论根据强关系还是弱关系来测量，银行都是处于最核心地位的行业。外展的国家关系网(主要由弱关系结合在一起)与内敛的区域集团网(由强关系结合在一起)的共存中体现了银行中心性。对它们来说，强关系具有明确的区域基础。内敛的区域集团是由金融企业和非金融企业的强关系产生的，但是，各个银行的强关系是强关系网的核心。例如，1962年的企业关系网就由一个非常大的关联成分(connected component)[18]、两个小的集团(每个集团有4个或5个企业)以及大量的企业对(pairs)和孤立企业构成。在大的关联成分中存在5个顶点及与之相连的聚类。在强关系网中的主导因素是一个区域聚类，该聚类围绕着"伊利诺斯大陆(Continental Illinois)"银行这个顶点和两个其他的芝加哥银行组织在一起，而这些银行与一个由11个中西部企业构成的群体关联在一起，这些企业又进一步与由132个企业构成的更大的集团关联在一起。在这个强关系网中，其他4个顶点是梅隆国民银行(Mellon National Bank)，摩根银行(J. P. Morgan)，信孚银行(Bankers Trust)以及加州联合银行(United California Bank)，与他们相联系

的各个聚类在规模上各不相同，从 4 个企业到 10 个企业不等。

弱关系把各个独立聚类连在一起，从而建构成一个巨大的国家网络，该网络恰恰处在上文所说的高度聚类的强区域关系网之上。比尔登等人认为，这种国家网反映了一种总体的向商业事件倾斜的取向，也反映出所有大公司都共有的利益趋同性。在非执行董事之间的连锁表明了这种共同性，并且产生了在国家层次上的整合性、一致性和互依性（也可参见 Useem，1984）。在该网络中，绝大多数公司都与一个大的单一成分相关联，绝大多数其余者都是独立的企业。银行当然还是最核心的行业，特别是那些起到"国家"而非"区域"作用的纽约银行。正是各个银行的非执行董事之间的连锁把整个国家网络胶合在一起[19]。

注 释

1 这一相关测度是通过公式：度数/($n-1$)来计算的，因为每一点最多可以与其他 $n-1$ 个点相连。

2 还可参见马斯登（Marsden，1982）。伯特发展了中间人观念，称之为"渔利者（*tertius gauden*）"，即从另外两派之间的冲突或分离中获益的第三派。安东尼斯（Anthonisse，1971）提出了一个称作"流（rush）"的测度，它与弗里曼的中心性观点高度相关。在 GRADAP 软件包中可以测量该测度。

3 作为一个比率，其取值范围为从 0 到 1，分值为 1 则表示一对点之间的连接完全依赖于 Y。

4 可见，中心度测度依赖于一组联立方程组的解。这一测量最初应用于比尔登等人（Bearden et al.，1975）的著述，其中 r_{ij}要根据两个公司之间连锁数量来测量。本章后文将讨论这种研究。

5 β 是一个附加在根据初始值测量出来的中心度上的权重，而不是一个乘数。

6 也可参见斯托克曼等学者（Stokman et al.，1985）对此观点的应用。

7 第 7 章中讨论的"位置"研究可以与斯托克曼和斯奈德兹的研究结合在一起，从而产生更可靠的区分。

8 在一个有向图中，"入"距离和"出"距离是不同的，因此，最大的列项将是"入-分离（in-separation）"，最大的行项将是"出-分离（out-separation）"。

9 注意：在这种情况下，点 B 与所有其他点都是等距的。在所有图中情况并不都是这样。

10 注意：由于不能定义出唯一一个绝对中心，因此通过克利斯托弗兹（Christofides）的程序定义出的真实的或想象的中心将有相同的数学特性。因此，在计算其他测度的时候就可以用绝对中心点。

11 如果一个图中确实存在唯一一个拥有最小离心率的点,那么该点就是这个图的绝对中心。如果如图 5.4 的社群图(4)所示,有两个具有同样的离心率的点,那么在这两点的中间就存在一个绝对中心。如果有两个以上的点具有最小离心率,则该图没有唯一的绝对中心。不幸的是,克利斯托弗兹算法在任何一个标准的社会网络分析软件中都不提供。

12 这些题外话对于社会网络分析的新手来说不是必需的,他们可能更喜欢直接从下一节开始。

13 在平面里,可以认为其面积为 $c(1.5)^2/3$ 或 $0.75c$。

14 圆周与直径之比就是一个常量,即圆周率 π。在一个图中,如已经在这里定义的那些概念一样,圆周与直径之比似乎不是一个常量。

我已经对在测量图的体积时的单位问题进行了注解。正如图论中所有的距离都用"线"来测量一样,这应该是测量的基础。因而,体积应当用"线的立方"来测量。

15 在公司网络中,强、弱关系之分后来在斯托克曼等学者(Stokman et al., 1985)的著作中得到了系统的阐述,这就像"首要(primary)"连锁与"松散(loose)"连锁之间的区别一样。

16 这一点是有争议的:在第二种情况下,权重也应该以作为接受者的董事会(recipient board)的规模为基础,因为给出权重就是为了试图测量连锁对于董事所在的董事会的重要性。对伯纳西茨测度的进一步思考可在米祖齐和邦廷(Mizruchi and Bunting, 1981),马里奥斯和琼斯(Mariolis and Jones, 1982)中找到。

17 这种聚类观点将在第 7 章中考察。

18 作为一个图的关联部分,"成分"这个观点将在下一章中给出更正式的定义。就目前讨论的目的来说,应当明确它的一般性意义。

19 有关这项研究与相关的研究工作的比较,可在斯科特(Scott, 1997)中找到。米祖齐和加拉斯科维兹(Mizruchi and Galaskiewicz, 1994)以及布拉斯和伯克哈特(Brass and Burckhardt, 1992)等也提供了一些有用的理论探讨。米祖齐(Mizruchi, 1992)将这项研究扩展到与之相关的公司的政治捐赠。

6 成分、核与派系

试图发现网络的各个“派系”以及网络可以分成多少个凝聚子群(cohesive sub-groups),这是投身于社会网络分析的学者们持续关注的问题之一。前文已经指出,早期霍桑实验和扬基城(Yankee City)研究者把“派系”思想看成是他们的核心理论发现。他们的论证是,人们之间的非正式关系把他们联络成为具有共同规范、价值、导向和亚文化的凝聚子群(sub-groupings),并且可以与“官方”或正式的社会结构相抗衡。他们认为,派系是人的身份和归属感的一个重要来源。他们广泛存在于日常词汇之中,如“我们组”、“后面的一组”等,人们正是运用这些词汇描述他们的社会世界。

一旦分析者试图把派系思想形式化,并且对派系的数量和凝聚力进行数学测量,人们才意识到,“派系”并不限于非正式的关系,还存在着政治派系和宗派、经济派系以及利益群体等。人们也发现,对“派系”这个简单的、明显的概念进行操作化测量的方法居然有多种:例如,可以把派系看成是相互联系的个体群,或者看成是密度高的群体。因此,各种关于子群的理论模型相继涌现,如把子群描述成为“派系(cliques)”、“聚类(clusters)”、“成分(components)”、“核(cores)”或者“圈子(circles)”等。这5个概念除了开头字母“c”相同以外,在其他方面很少有共同之处。本章将讨论它们的各种理论基础,尽管我把“聚类分析”问题挪到下一章讨论。

所有这些对群体结构进行测量的起始点都涉及“子图(sub-graph)”这个概念。一个子图指的是从一个网络图中选择出来的任何点和连接这些点的线构成的集合。图的任何部分都可以选出来进行子群分析,尽管不是所有这些标准都能在研究中得到实质性应用。例如,一个随机的点样本可以被看成是一个子图,从而可以考察其结构特征。但是,一般来说,一个随

机子图很可能不对应于一个有意义的社会群体。一种比较有用的区分子图的研究就是利用一些标准(如性别)把网络的各个成员分开,从而可以分析两个独立子图:男子图和女子图。任何这种选择都依赖于研究者的理论关注点和经验关切点。其总目标仅仅是界定一个有意义的能动者类别,探讨它们的网络形式的独特模式。因此,从这种观点看,子图的确定与最初对各个图本身进行界定是一样的。前几章讨论的边界和抽样问题同样也与此处的问题有关,这里不涉及新问题(参见 Frank,1978b)。

派系以及类似的分析通常采用另一种研究方式来探讨子图问题。其目的是考察整个图本身的结构特征,以便发现该图可以分成几个"自然存在的"子图。从这一观点出发,一个子图必然具有某种根据图论数学原则得出来的特征:点的关联性、关系的紧密性等。它是一个对于某一独特特征来说达到**最大**的子图:这种独特特征如果不消失的话,它将是图中存在的最大子图。至于特定特征的选择,则完全依赖于研究者自己的决定,即研究者要决定给某个特定的数学标准以有意义的、有用的社会学解释。不幸的是,这一做法很少是明显的,并且绝大多数研究者都假设,不管在社会网络分析软件中应用什么样的数学程序,这些程序都**必须**是有用的社会学测度。本章的目的就是揭示在各种现有的程序中存在哪些数学假设。了解这些假设之后,研究者才可以就他们的特定研究作出有根据的决定。

成分、循环和结群

在各种子图概念中,最简单的为**成分**(component),它的正式定义是:"最大关联的子图"。与图一样,在子图中,当其所有点都通过各种途径相连的时候,子图就是"关联"的:一个关联子图内的所有点都可以通过一条线或者多条线相互连接,但是它们与子图外的点无关联。在一个成分中,所有点都通过"途径"连在一起,但是其中任何线都不指向成分外的点。当关联子图达到最大的时候,不破坏关联性就不可能加入新成员。例如,不能把一些孤立点加到一个现存的成分之中,因为它们与该成分的任何成员之间都没有关联。因此,确定一个成分的边界需要通过追溯来自其成员的各条途径,以检验其关联性。

用来区分成分的计算机算法可以开始于一个随机选择出来的点,并且追溯与之直接相连的所有其他点。对找出来的这些点也可以应用同样的分析程序,用这种"滚雪球"方法,成分的规模因而将逐渐增大。当不存在

其他可以加入到该成分中的点的时候，成分的全部成员就确定下来了。如果成分之外还有其他点，可以针对这些点进行同样的分析，以便发现图中是否可以分析出其他成分。

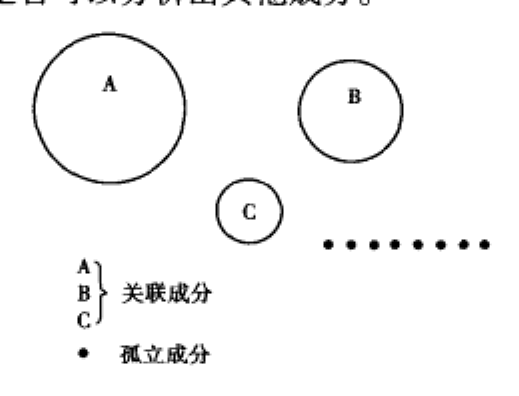


图6.1　网络中的成分

这样看来，成分是一个点集，这些点通过连续的关系链连在一起。可以追溯连接这些点的途径，直到发现成分的边界。当然，一个"关联图"仅仅由一个单一成分构成。其他图主要由一个或多个独立的成分以及一系列孤立点构成（参见图6.1）。这种观念易于进行社会学解释。从原则上讲，一个成分的各个成员都可以相互交往，无论这种交往是直接的，还是通过一条中介链。另一方面，孤立点则无此机会。因此，在一个图中发现的各个成分的模式——其数目和规模——可以看成是标志了其成员进行交往的机会和限制性，或者表征了相应的网络中的资源流动。因此，就此来说，它们体现了早期的场域理论中的"拓扑式区域(topological regions)"背后的观念。因此，对一个网络进行结构描述的第一步便是确定其成分的数目和规模。

探查一个图中各个成分的最简单算法是找出所有可能的途径以便找到两点之间的捷径。这些捷径的长度最小值为1（直接联系），最大值为$n-1$。例如，在一个规模为100的图中，途径的最大长度可以达到99。然而，在大型图中，一个成分的最长捷径——其"直径"——一般要远小于此[1]。但是，由于在确定成分的界限之前一般不知道成分的直径，所以这种算法必须考察到所有的途径，一直到最大值$n-1$，这样才可以找出直径。

由于该程序用时长，效率低，因而对于大多数计算目标来说不适用。因此，社会网络软件包一般运用另一种替代程序。通过构建"生成树(spanning trees)"，利用从所选择出来的各个点回溯(backtracking)的方法发现"成分"。该算法先找出与某个开始点相连的任意一点，再找出与这个任意点相连的任何点。如此反复，直到找不出另外的关联为止。然后，该算法沿着所发现的链返回，直到它能够与一个新点建立联系。重复此类回溯方法，就可以很高效地找到一个成分的边界，并且该程序可以继续考察其他点，以便发现其他成分。

无论对于有向图还是无向图来说，都可以找出其成分，但是这两种情

形有很大区别。在有向图的情形中，可以确定两类成分：强成分（strong component）和弱成分（weak component）。前者构成途径的各条线排列在一条持续链中，没有任何方向上的改变。任何不满足此标准的途径都不加以考虑。之所以这样限定，原因在于我们假设一条线的方向表明某种资源或者工具（如资金、权力或者信息等）可能的流动方向。只有在途径中的各条线不改变方向的时候，资金或者信息的流动性才不能被打破。这样看来，强成分代表的是一系列行动者，在这些行动者中，上述资源可以轻松自由地流动。

对有向线也可以有另外一种弱解释。我们可以假设，只要存在关系，不管其方向如何，都有可能出现交往。按照这一观点，可以从图中的各个半途径（semi-paths）中确定出成分。由于不考虑构成途径的各条线的方向，仅考虑关系的有无，因此，用这种方法分析得到的有向图中的成分称之为弱成分。

当然，在无向图中不存在成分的强弱之分。此时研究者着眼的往往是简单的成分（simple components）：由于所有线都没有方向，所有途径都由可接受的关系构成。原则上说，区分一个无向图中的简单成分的计算机程序与区分有向图中的弱成分的程序是一致的。只有当线的方向需要单独加以考虑的时候，算法才有区别。

成分分析的结果是把图看成是由一个或者多个（简单的、弱的或者强的）成分和一系列孤立点构成的。紧密的图一般由一个大的单个成分占据主导地位，特别是当分析关注的是简单的或者弱成分的时候更是如此。为了获得一种更加缜密的分析，一般需要探究成分的内在结构。

埃弗里特扩展了成分观念，目的是更加细密地分析致密网络之结构。他的研究（埃弗里特，1982，1983a，1983b，1984）建立在一个他所说的“块（block）”这个图论概念基础之上。“块”一词是极为混乱的，它在社会网络分析中也有各种各样差异极大的所指。为了避免混淆，我建议在术语上要有革新。由于在下文中即将明确指出的原因，我在这里用“环成分[2]（cyclic component）”代替埃弗里特的“块”概念。

“环成分”概念依赖于“循环（cycle）”一词。一个**循环**就是一个途径，只不过它返回到其初始点，并且与一个途径类似，其长度也可以是任意的。一个图中的循环可以用其长度来描述，如3-环、4-环等。就其最一般的形式来说，图论专家可以确定埃弗里特所说的*k*-环，*k*可以是任何特定环的长度。分析环的第一步往往是决定要考察的环最长是多少。任何比它长的环都忽略不记。例如，如果选择最大环长为4，那么，图6.2中的（1）则包含4个长度为4的环（ABCDA、BCDAB、CDABC和DABCD）和6个长度为3的

环(ABDA、BDAB、DABD、BCDB、CDBC 和 DBCD)[3]。如果要求最大环长为3,那么只有较短的环保留下来,此时的点 A 和 C 就不被任何环连在一起。埃弗里特接着引入"桥线(bridge)"一词,把它定义为这样一条线,即它本身不处于环中,但是却把两个或者更多的环连在一起[4]。例如,就最大环长为4而言,图6.2中(2)包含的桥线是BE。

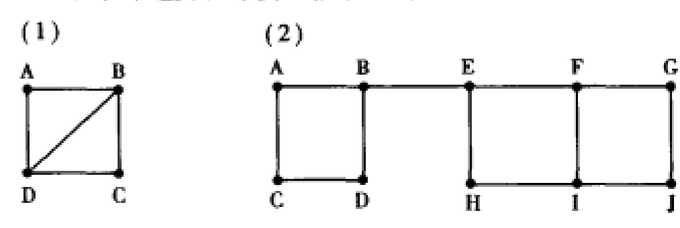


图6.2 环成分

一个"环成分"可以定义为一系列互相交叉的环构成的集合,这些环由它们共有的线或点连在一起。这样,尽管一个图的多个独立环成分之间通过一条或者多条桥线连在一起,但是它们之间却不重叠。例如,图6.2中(2)本身不是一个环成分,然而它却包含环成分{A、B、C、D}和{E、F、G、H、I、J}。后一个点集包含线FI,它是EFIHE和FGJIF这两个环的公共线。因此,可以看出,一个环成分是由一个相互交叉的环链构成的,相交处是重叠的环共有的线或点[5]。把一个图中所有那些特定环长的桥线(称为 *k* 桥线,*k*-bridges)从图中移开,就可以确定各个环成分是什么。剩下的点集就是环成分。

对简单成分、弱成分和强成分的分析得到的只不过是一些成分和孤立点,而对环成分的分析得到的常常是比较复杂的结果。这是因为,环成分与那些本身不是环成分成员的各种点连在一起。埃弗里特(Everett,1982)指出,这些联系要素可以是如下5种之一:

1. **环成分**。

2. **悬挂点**(hangers)。这是一些与环成分相连之点,但它们本身不在环上。这些点仅仅"挂"在环成分上。

3. **桥点**。这些点处在两个或者多个环成分之间,起到中介者(intermediaries)或者"摆动者(waverers)"的作用,但它们不是任意环成分的成员。这样,一个桥点"挂在"两个或者多个环成分上。

4. **孤立树**(isolated trees)。这是一些点链(包括二人链(dyads)),它们与任何环成分都无关联。这些树的成员之间以非环(non-cyclic)的方式相连[6]。

5. **孤立点**。是一个度数为零，与任何点都不相连的点。

有时候很难给长的关联途径以有实质意义的社会学解释。当长环把众多点连在一起的时候，问题就更加明显。例如，关联图有一种趋势，即它是由一个大的循环成分构成的。出于多方面考虑，埃弗里特认为，在现实情形中应该把分析限定在相对较短的环上，如3-环或者4-环。如果分析3-环，那么仅仅关注由三人组（triads）构成的环成分就可以了，并且可以给出很多具有实质意义的解释。分析4-环，则需要考察三人组或者矩形（rectangle）关系。尤其重要的是，如果研究者希望运用长度超过4的环，那么他赋予这种数学结构的社会学解释必须既清楚又有意义。

对于有向图来说也可以分析其环成分。最简单的做法就是不考虑线的方向。这种分析的基础是图中的"半途径（semi-paths）"，分析出来的是"半环（semi-cycles）"。在这些环中都不考虑线的方向。当然，这样做会失去一些信息，但是可以得到**弱环成分**（weak cyclic components）。为了分析**强环成分**（strong cyclic components），方向性信息必须保留。事实上，埃弗里特建议，这种分析也必须包含某些半-环。在一个有向环（directed cycle）中，方向沿着所有线依次传导下去。例如，在图6.3中，ABCA是一个有向环。而途径ABDA则包含一个A和D之间的反方向，所以它仅仅是一个半环。当那些不在有向环上的点是由两条或者更多有向途径连在一起的时候，埃弗里特把这种半环定义为"可接受的半环"。这样，点A和D不通过一个有向环连在一起，但它们是通过有向途径ABD和AD连在一起的。因此，ABDA是一个可接受的半环。

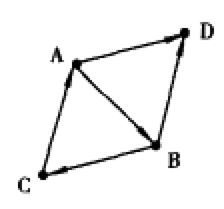


图6.3 环和半环

因此，在确定强环成分的时候，计算机算法必须搜索一个图中的有向环和可接受的半环。利用这种程序，一个有向图中的所有环都可分为有向环、可接受的环和不可接受的环三类，并且对强成分的分析仅仅需要考虑前两种环即可。仅用这些环就可以确定图的强环成分，同时也可能根据连接这些点的线的方向，对"在悬点（hangers-on）"和"未悬点（hangers-off）"进行区分。所谓"**在悬点**"指的是这样一些悬点，即连着它的线指向强环成分的某一成员。反之，所谓"**未悬点**"指的是这样一些悬点，即存在一条从成分的一个成员出发并指向该点的线[7]。

还有一种探究成分的内在结构的方法，即考虑是否存在一些把各个成分联络在一起并且具有轴心作用的关键点。与埃弗里特一样，哈吉和哈拉里（Hage and Harary，1983）也通过所谓的"块"概念来研究此问题。然而，他们把"块"看成是在没有切割点[8]（cut-point）的简单成分（或者在一个有

向图的弱成分)中的一些子图。所谓**切割点**指的是这样的点,即如果去掉它,就会增加成分的数目,没有它,子图就会分为两个或多个独立的子群,并且各个子群之间无关联。例如,在图6.4的成分(1)中,点B就是一个切割点,因为一旦没有B,(1)就变成了两个无关联的成分,即变为(2)了。其他点都不是切割点。

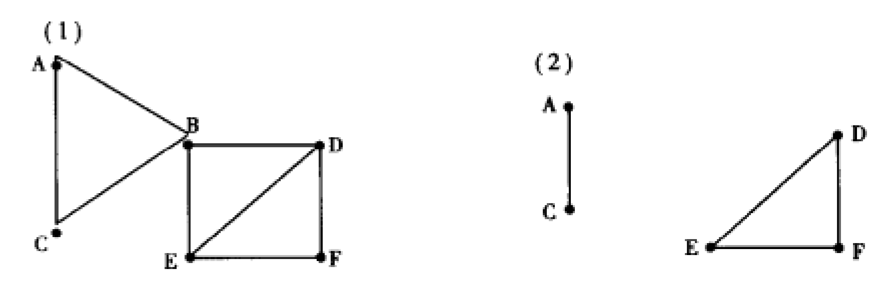


图6.4 结群和切割点

这样看来,切割点是在构成一个成分的各个要素之间起到中枢作用的关键点。这些要素(elements)和切割点便是哈吉和哈拉里所描述的"块"。另外,这里还要声明的是,为了避免在"块"概念的各种不同应用方面出现混淆,我在下文将不用"块"这个词,而用"结群(knot)"这个描述性词语代替它。如此看来,图6.4中的成分便由两个结群{A,B,C}和{B,D,E,F}构成。因此,一个图中的各个切割点将成为大量结群的成员,切割点成为各个结群[9]之间的交叉点。

相对而言,比较容易给切割点以具有实质意义的社会学解释。例如,可以认为切割点代表的行动者处于局部中心地位。例如,哈吉和哈拉里(Hage and Harary,1983)曾经指出,结群(他们用"块"一词)可以看成是一个网络中最有效的交往或者交换系统(该观点的具体应用可参见Hage and Harary,1991,1998)。由于结群不包含切割点,其各个成员之间的交往和交换行为因而不依赖于任何单个成员。在结群的所有点之间总是存在替代的交往途径,这使得其形成的网络既有弹性又不分层化。

成分的轮廓

到目前为止,我们已经探讨了区分各种成分的程序,对分析这些成分(结群、切割点)构成要素的方案进行了述评,也述评了外在于成分的一些

要素(悬挂点、桥线、树和孤立点)。在以下两节中,我将较系统地考察成分的内在结构问题。在本节中,我将探讨如何通过找出成分的"核"而画出其轮廓(contour),下一节将分析用来构建成分的"派系"和"环(circles)"。

在第2章中我已经指出,扬基城(Yankee city)的研究者试图找出他们所说的"派系"的"核心成员"和"边缘成员"。这个步骤也可用于对成分内在结构的分析。通过一种常常被称为成分"嵌套(nesting)"的程序(这在第3章已有简单的分析)[10],就可以揭示成分的轮廓。对成分进行层层嵌套分析涉及比较强的临界(cut-off)标准,以便在分析的每一步都能画出成分的边界。当把分析的结果组合为一个图形的时候,得到的就是一系列有公共圆心的点集。在嵌套分析中,基本图形是一个轮廓地图或者一系列俄罗斯套娃①(Russian dolls),每个成分都被"嵌套"在更大的成分之中。一个成分被可视化为以一些紧密凝聚的点为核心,随着凝聚力或者紧密度标准越来越弱,其边界也逐渐向外扩散,包括越来越多的点。在联系最弱的层次上,所有的关联点都包含在一个单一的成分之中。

图6.5展示了一个简单的嵌套案例。集合A的点之间联系最密切,它们构成了成分的核心。确定集合B的边界的标准就弱一些,所以它包含集合A的所有点和另外一些联系较弱的点。最后,用来确定集合C的边界的关联标准是最弱的,所以包含所有的点。第二个成分中,集合D、E、F等也可以给出同样的解释。这样,一个图中的每个成分都可以解构为各个核心要素,从而可以画出该图的轮廓图。

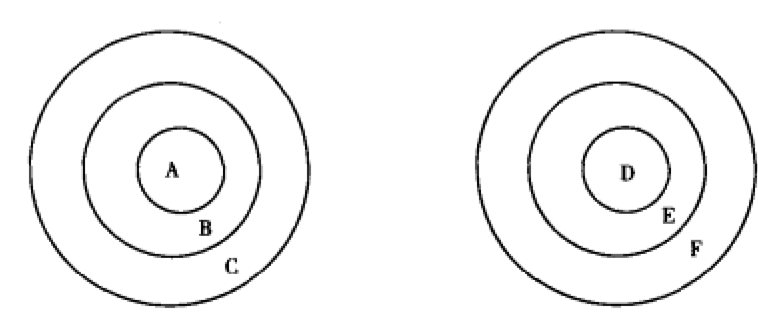


图6.5 嵌套的成分

检验成分的各种算法把所有的关系都看成是二元数据,即简单地看关系的有无。因此,对一个多值图的分析必须把真实值转换为二元值,即1

① 全文应该为Russian nesting dolls,是一种最受欢迎的著名俄罗斯纪念品,其中的多个玩具层层嵌套在一起——译注。

或者0。这便需要对多值图矩阵中的各项进行如下"切割"或者"二分法"操作[11]:高于或者低于某个特定临界值的各项被转换为二值数据:高于临界值的各项赋予"1",低于此值的各项赋予"0"。然后就可以运用这些二元数据寻找成分了。例如,一个多值邻接矩阵的各项可能表达线的多元度,这个矩阵就可以根据逐渐增强的密度标准进行"切分"。通过考察根据每类临界值确定下来的成分,研究者就可以构建如图6.5所示的嵌套图轮廓。成分的边界具有共同的中心,图形也表示了高密度的"峰值(peaks)"和低密度的"平坦"。

人们提出两类相互替代的嵌套方法:一类根据点的度数作为测量凝聚力的标准;另一类根据线的多元性作为测量紧密度的标准。以度数为基础(degree-based)的测量确定的是"*k*-核(*k*-cores)",基于多元性的测量确定的则是"*m*-核(*m*-cores)[12]"。

赛德曼(Seidman,1983)指出,对成分结构的研究可以运用最小度标准,以便区分高、低凝聚力的领域。他论证道,对一个图的"*k*-核"结构分析是对密度测度的一个重要补充,笔者已经指出了密度测度不能抓住图结构的许多整体性质。一个***k*-核**是一个最大子图,其中的每个点都至少与其他*k*个点连接:*k*-核中的每个点的度数都至少[13]为*k*。这样,一个简单的成分就是一个"1*k*-核"。其中所有点都相连,因而其度数至少为1。为了确定一个"2*k*-核",需要忽略所有度数为1的点,进而考察剩余各个点之间的关联结构。2*k*-核由那些度数为2的剩余关联点组成。同理,确定一个3*k*-核则要去掉度数为2或1的点,依此类推。

图6.6表示的是一个3*k*-核图。在这个子图中,每个点的度数至少为3。尽管有两个点(B和J)的度数是4,但是该图不存在4*k*-核,因为一个*k*-核必须至少有$k+1$个成员。

如此看来,一个*k*-核便是在整个图中的一个凝聚力相对较高的区域。但它不一定是最大的凝聚子图,因为有可能存在一些相互之间联系松散,但却有很高凝聚力的区域。例如,在图6.6中,凝聚区域{E,F,G,H,J,K,L}和{A,B,C,D,E,I}是通过弱关系线IJ和CE连在一起的。这样,这些*k*-核就构成了成分的各个域,在该成分中可发现具有凝聚力的子群(如果存在的话)[14]。

赛德曼也指出了如何估计一个网络的总体分裂性(fragmentation),这就需要根据他所说的**核塌缩序列**(core collapse sequence)。一个*k*-核中的点可以分为两个集合:在$k+1$核中的点和不在该核中的点。赛德曼把后一群体称为*k*-剩余集合(*k*-remainder)。在任何核中,剩余集合都是由那些当*k*增加1后在分析时将会"消失"的点组成。正是这些当*k*增加时的关

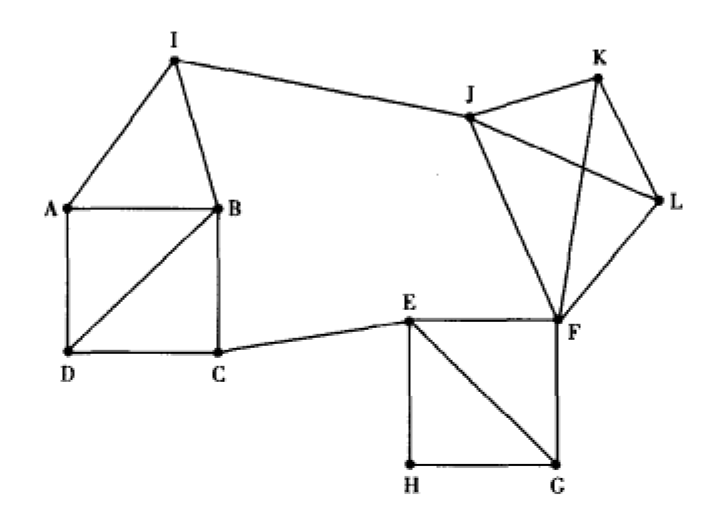
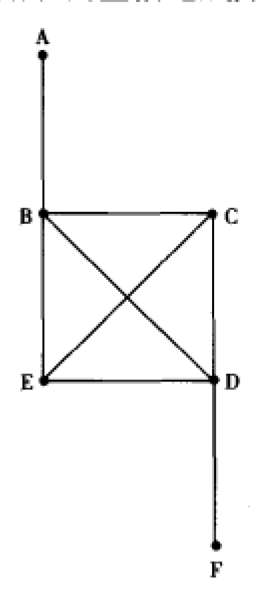


图 6.6　一个 3k-核图

联较小点的消失才导致了“塌缩(collapse)”。赛德曼提出，每当 k 增加一个单位，从核中消失的点所占比例可以排列为一个向量(即一行简单的数值)，可用该向量描述成分内部的局部密度结构[15]。



k 值	剩余点	剩余点所占比例
0	0	0
1	2	0.3
2	0	0
3	4	0.6
4	无	无

图 6.7　k-核的塌缩

这一点可以在图 6.7 中得到展示。在该图中，所有点都连在一起，所以 k 从 0 增加到 1 时不会引起点的消失。当 $k=1$ 的时候，所有点都包含在一个核内，但是有两点是剩余的(点 A 和 F)。当 $k=2$ 时，剩下点 B、C、D、

E,每个点的度数都不小于2。事实上,由于这些点的度数都是3,并且都相互关联,所以当 $k=2$ 时就无剩余点了。然而,当 k 增加到3的时候,剩余4个点,因为全部点在 k 增加到4的时候都消失。如果用一个向量来表示从 $k=0$ 开始出现的剩余点,可以得到如下核塌缩序列:(0,0.3,0,0.6)。

核塌缩系列汇总了成分的密集性(clumpiness)。赛德曼认为,如果核塌缩是缓慢的、逐渐的,则表示网络结构在总体上具有一致性。否则,一个不规则的取值序列(如图6.6所示)则说明存在着相对来说比较紧密的区域,该区域被比较多的边缘点包围着。如果向量中的值从0持续增加到比较高的 k,这表明成分内的结构具有一致性;如果在较低的 k 值出现以后持续出现了0值,则表明存在多个的高密度。

k-核以点的度数为基础。相比之下,m-核则基于线的多元度(multiplicities)。m-核这个概念描述的是由GRADAP研究小组[16]讨论的初始嵌套成分。一个 m-核可以定义为一个最大子图,其中每条线的多元度都不小于 m。一个 m-核是由多条具有特定多元度的线连成的一系列点链。与 k-核一样,一个简单成分也是一个1m-核,因为其中所有点都由多元度至少为1的线连在一起。在一个2m-核中,就忽略多元度为1的线,根据剩余的线来确定成分。一个3m-核也类似,忽略所有多元度为2和1的线。图6.8表示的是一个简单的3m-核,其中所有点都由多元度至少为3的途径相连,各个点与核之外的弱关联不被考虑。由于点B和C由一条多元度为4的线相连,它们形成了一个2-成员的4m-核。正是各个核的相互嵌套才揭示了网络的总体形状[17]。

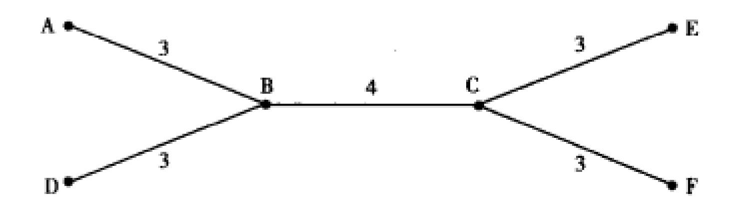


图6.8 一个3m-核

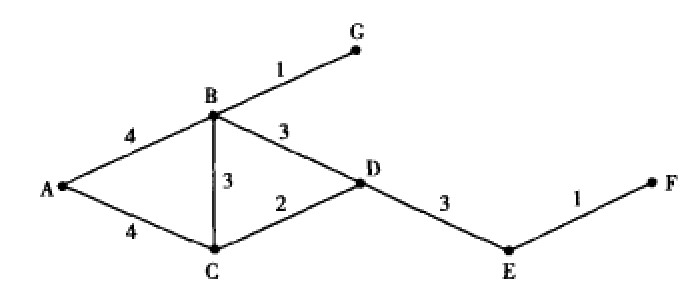
赛德曼关于核塌缩系列的思想也可以扩展到 m-核;当然,这一思想距能实际应用还有很大的距离。这一思想可以用图6.9来解释。随着 m 值的增加,线被逐渐撤除。在每一个 m 水平上的剩余的线数是当 m 增加到 $m+1$时将会消失的点数。当 m 由1增加到2的时候,有2个点将会消失,但是其他点不会再消失,直到 m 增加到4的时候。如果 m 增加到5,所有点都将消失,因为图中点的最高多元度是4。这样,m-核塌缩系列是(0,

0.28,0,0.28,0.43)。

为了完成本部分论述，有必要考虑嵌套分析与环成分之间的关系。当然，环成分可以在多值图中区分出来，区分方法是运用一个适当的"切割"值。运用不同的标准，即可达到对嵌套成分的分析——在这种情况下是对嵌套环成分的分析[18]。

总之，对简单成分基本思想的各种推广为分析网络裂变程度提供了一系列有力的概念工具。

它们通过强调网络的总体形状补充了密度测量，有助于克服密度测量的许多局限性。如果针对规模可比的网络的各种整体结构进行全方位的比较研究，那么将涉及测量网络总体密度、内涵度、成分的数量和规模、成分的密度、成分的嵌套结构与核塌缩系列等。



m 值	m-核	剩余点数目	剩余点所占比例
0	{A,B,C,D,E,F,G}	0	0
1	{A,B,C,D,E,F,G}	2	0.28
2	{A,B,C,D,E}	0	0
3	{A,B,C,D,E}	2	0.28
4	{A,B,C}	3	0.43
5		无	

图 6.9　一个 m-核的塌缩

派系及其交叉

到目前为止，本章讨论的各种概念已经指向了派系思想的形式化，早期社会网络学者们在霍桑工厂和扬基城的研究中已经发现派系的存在。

但是我还没有讨论社会计量学意义上的派系概念本身，这个概念是人们在讨论图论的社会学应用之时提出来的。对“派系”有各种相互对抗性的应用，但是，大多数人都同意把派系的本质含义看成是“最大的完全子图（maximal complete sub-graph）”（Harary，1969；Luce and Perry，1949）。也就是说，一个**派系**就是一个点的子集（seb-set），其中任何一对点都由一条线直接相连，并且该派系不被其他任何派系所包含[19]。如图6.10所示，一个3-成员的派系包含三条线，一个4-成员的派系有6条线，5-成员的派系有10条线，依次类推[20]。一个“成分”是“最大”的关联图（所有的点都通过途径相连），而一个“派系”也是最大的完全关联图（所有点都相互邻接）。

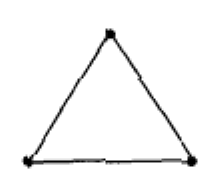
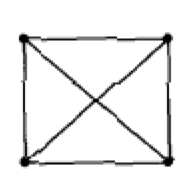
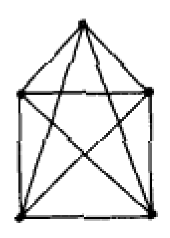




图6.10　不同规模的派系

多瑞安（Doreian，1979：51-2）指出了派系的一些形式特征。其最基本的特征是，所有派系都是点的最大子集，也就是说，其中每一点与其他点都处于直接的互惠关系之中。根据定义，在一个无向图中，所有点之间的关系都是相互的，所以派系分析程序要用到图中全部线。然而，在有向图中情况则有所区别：其矩阵是不对称的，只有双向的关系才予以考虑。因而，在有向图中，网络分析区分出来的派系叫做**强派系**（strong cliques）。另一方面，如果不考虑有向图的方向，仅仅关注关系的有无，把所有的线看成是双向的关系，那么这种分析的结果就是**弱派系**[21]（weak cliques）。

最大完全子图这个概念对于社会网络来说是相当严格的，这种紧密联系的群体在实际中很少见。因此，有学者对此概念进行了一系列推广[22]。最早给出的推广是n-派系（n-clique）这个概念，这个概念与人们日常对派系的理解最接近。在此概念中，n指的是派系成员之间联络的最长途径之长。这样，一个1-派系就是最大完备子图本身，因为其中所有点都直接相连，距离都是1。一个2-派系则是这样的一个派系，即其成员或者直接相连（距离为1），或者通过一个共同邻点间接相连（距离为2）。当然，在分析中所使用的n的大小要由研究者自己决定，n越大，对派系成员限制的标准就越松散（参见图6.11）。例如，一个3-派系就是一个比2-派系松散的群体。n可能取的最大值要比图中点的总数少1。然而，在实践中，大多数大的关

联图都可加入到一个 n-派系之中，并且这里的 n 可以远远小于最大值。

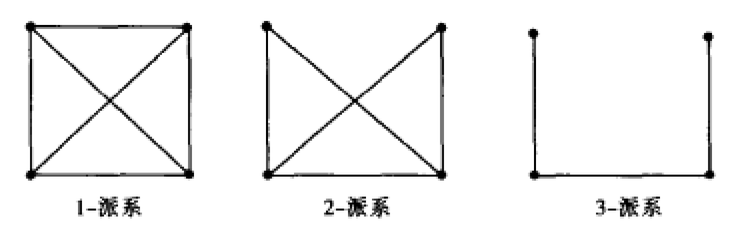


图 6.11 规模为 4 的 n-派系

n-派系可以通过相对简单的矩阵操作方法进行分析，在许多数据表(spreatsheet)程序或者社会网络分析专业软件之中都有这种方法。例如，对邻接矩阵本身进行操作，就可以产生一个途径距离矩阵。邻接矩阵的平方表示的是所有距离为 2 的关系，邻接矩阵的三次方给出所有距离为 3 的关系，依次类推。然而，矩阵操作法是一种效率低的派系探查方法，大多数专业网络分析程序利用另外一种程序，即在分析成分时所使用的回溯程序的一种变形。由于这种程序简单易行，无向图的派系分析因而成为大多数社会网络分析程序的特点[23]。我们也可以分析多值图中的 n-派系，其方法与前一部分所讨论的一样，也要利用一种切割标准。对于 n 的每一个层次值来说，这种分析就会产生一系列嵌套派系：嵌套 2-派系，嵌套 3-派系等。

n-派系思想在应用方面有两个严重局限。首先并且也是最主要的局限是，当 n 大于 2 的时候，很难给它以社会学解释。距离为 2 的关系可以直接解释为那些有共同邻点的人之间的关系，该邻点可以起到中间人或者掮客的作用等。然而，长度大于 2 的途径则包括比较疏远的弱关系。尽管比较长的弱关系链可能对网络的总体结构来说非常重要，如格拉诺维特和小世界分析者(small world analysts)所指出的那样，但是，它们对于派系的定义的适用性是什么，这一点根本就不是很清晰。一谈到派系，就要求有相对紧密的关系。因此，很难为 n 大于 2 的 n-派系的区分进行辩护。

派系思想在应用方面的另一个局限是存在如下事实：n-派系的一些途径中的中介点本身可能不是派系的成员。例如，图 6.12 中(1)的点 A、B、C、D、E 形成一个 2-派系，但是，连接 D 和 E 的长度为 2 的途径要经过点 F，而点 F 却不是该 2-派系的成员。派系的“直径”(即最长距离之长度)可能比用来定义 n-派系的 n 长。这样，点集{A、B、C、D、E}由一个 2-派系构成，但是它的直径却是 3。阿尔巴(Alba，1973，1982)和莫肯(Mokken，1974)二人都提出了这个问题，并对 n-派系思想进行了许多推广。莫肯认为，应该再提出一个有用的概念，把 n-派系的直径限定为 n。也就是说，假设研究者

接受距离为 2 来确定派系的成员，同时也要求派系的直径不大于 2。他称这一概念为 n-宗派（n-clan）。图 6.12 中的（2）和（3）是 2-宗派[24]，而（1）不是。

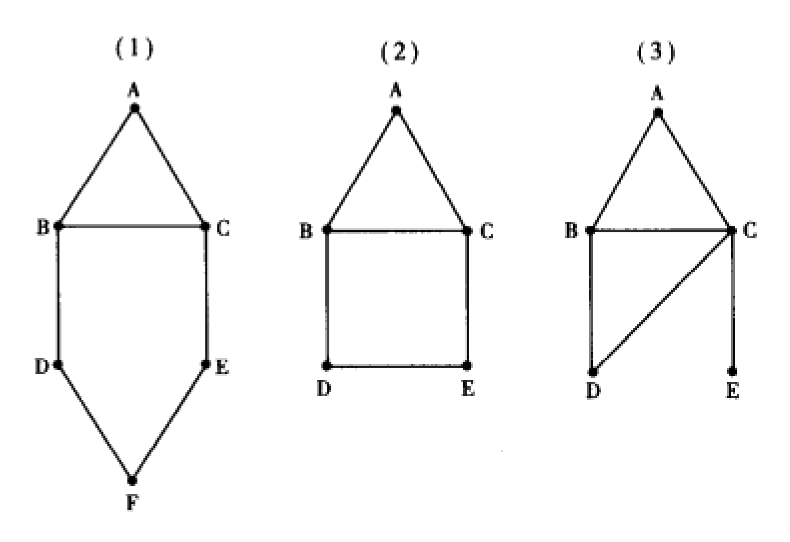
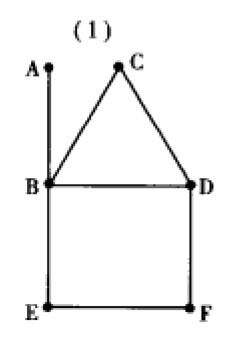


图 6.12 子图和 2-派系

针对派系基本思想的另一个推广是赛德曼和福斯特（Seidman and Foster, 1978）提出来的 k-丛（k-plex）观念。n-派系概念涉及增加用来界定派系的途径长度，而 k-丛概念关注减少每个点所指向的其他关联点的个数。这样，一个 k-丛中的点以距离 1 相连，但是，并不是所有点都相互关联。一个 k-丛就是满足如下条件的一系列点，即其中每个点都与除了 k 个点之外的其他点直接相连[25]。这样，如果 $k=1$，那么一个 1-丛就等于 1-派系，所以它是一个最大的完全子图。1-丛中的每一个成员都与其他 n-1 个点相连。当 $k=2$ 的时候，其中所有点都至少与 n-2 个其他点相连，但是 2-丛可以不是 2-派系。在图 6.13 中，（1）是一个 3-派系，因为所有对点之间的距离都不大于 3。然而，它却不是一个 3-丛，因为点 A、C、E、F 中的每一个连接的其他成员都少于 3 个。图（2）则既是一个 3-派系，也是一个 3-丛[26]。

在分析 k-丛时，需要考虑的一个重要问题是，研究者如何确定一个丛的可以接受的最小规模。尤其需要考虑的是，k 值越大，导致应该可以接受的 k-丛规模的临界值（cut-off threshold）越高。当 k 取值较低的时候，k-丛可以是相对较小的，但是，k 值增大将会产生无用的结果，除非增加可以接受的 k-丛的最小规模值。其原因在于，k 值高的小子图只能是凝聚力较小的子图。如果进行约略的估计，那么一个可以接受的 k-丛的最小规模应该是



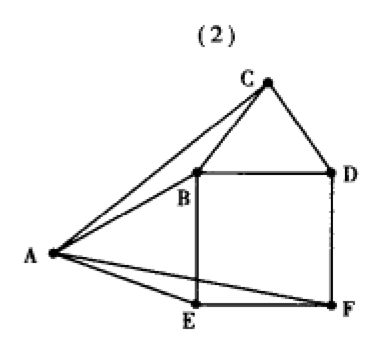


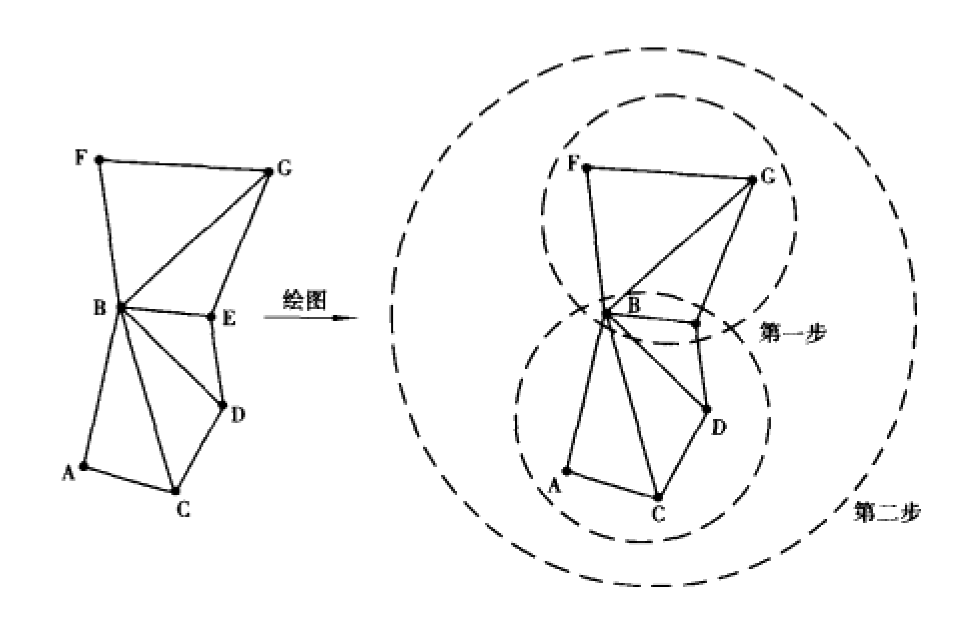
图 6.13 3-派系和 3-丛

$k+2$。然而，k-丛这个概念（被认为是基本派系思想的推广）似乎比 n-派系更能体现凝聚力思想，当 n 的取值比 2 大的时候更是如此[27]。与 n-派系和成分的情况类似，k-丛的基本思想也可扩展到对多值图的分析，这还要用切分标准（slicing criterion）来分析"嵌套的 k-丛（nested k-plexes）"。

除了最小图之外，在任何图中，构成图的各种 k-派系和 k-丛之间都存在大量的重叠。派系分析（既包括 n-派系，也包括 k-丛）趋于产生一长串重叠的派系，这些结果难以解释。一个相对密集的网络常常由大量互相重叠的派系构成，许多点都是大量派系的成员。例如，对于一个包含 20 个点的密集网络来说，它可能包含大约 2 000 个互相重叠的派系。在这些情况下，派系之间重叠的密度可能比派系本身的构成更有意义。因此，阿尔巴（Alba，1982）指出，社会网络分析者应该应用那些明确体现出这种重叠事实的概念。在卡杜申（Kadushin）和摩尔（Moore）（Alba and Kadushin，1976；Alba and Moore，1978；Kadushin，1966，1968）的研究基础上，他认为，"社会圈（social circle）"这个概念可用来抓住社会网络的重要结构特征。

这个观念是卡杜申根据齐美尔（Simmel，1908）的最初洞见提出来的，后者最早提出了"社会圈交叉"的重要性。一个社会圈的凝聚性不建立在其成员的'面对面'的接触之上，而是建立在把人们连在一起的相对较短的间接关系链的基础上。社会圈从互动中"出现"，并且可能不被参与者注意到，因为其边界仅仅是由这些间接关系的分支松散地界定的。

阿尔巴的贡献在于，他用社会计量学的术语对"圈"这个概念进行了形式化处理，方法是建立这个概念与其他图论概念之间的联系。他的基本论证是，如果重叠的派系拥有一定比例的公共成员，那么它们可以组合在一起成为一个"圈"。阿尔巴建议，最恰当的程序就是应用某种"滚雪球"方法，这样，派系就会集结成为越来越大的松散的圈。在分析圈的时候，第一步是确定规模为 3 的 1-派系（三人组），然后把它们合并到一个圈中，其中



1-派系：{A、B、C}{B、C、D}{B、D、E}{B、F、G}{B、G、E}
第一步产生的圈：{A、B、C、D、E}{B、F、G、E}
第二步产生的圈：{A、B、C、D、E、F、G}

图6.14 互相交叉的社会圈

所有的派系都只有一个成员不同。换句话说，在第一步中，确定圈的标准是，如果派系的三分之二的成员完全相同，就把这些派系合并为一个圈。这样，第一步分析的结果是一个或者多个圈，以及一系列独立的派系和孤立点。第二步，剩余的派系可能合并到另外一些圈当中，这些圈之间的重叠层次较低。阿尔巴建议，在这一步中有三分之一的重叠就可以。这种组合的结果是产生一个大圈或者一系列被联系不很紧密的派系和点包围的小圈。图6.14表示了一个简化的社会圈分析。第一步分析出来两个圈，但是它们在第二步合并在一个圈中。与许多图论程序一样，有很重要的一点需要指出，即选择什么样的重叠水平作为标准，这是没有定论的。阿尔巴建议的重叠水平是在常识性的数学背景中选择使用的，研究者必须清醒地知道，他的建议在特定应用中是否有意义。

因此，对圈的测量把派系的重叠程度看成是派系之间距离的一种测度。在这种分析程序中，用来分析派系（如 *n*-派系或者 *k*-丛）的特殊方法几乎用不上，因为在组合的过程中这些程序之间的细微差异迅速消失。在实际中，最后结果是组合为很多圈，这很少受到最初用到的派系检测方法的影响[28]。

成分和引文圈

许多研究领域都涉及社会网络观念，科学社会学是其中之一。克兰(Crane，1972)的《无形学院》①(*invisible college*)是最早利用科学家之间交流网络观念的研究成果之一，她用这种观念来解释科学知识的增长途径。克兰的著作使用了问卷，获得了关于农村社会学家之间的交流模式和影响方面的信息。她分析了诸如合作发表论文，在专业研究领域提供建议这样的现象。她的关注点是勾勒出在专业研究领域中合作者的无形学院的规模和重要性，但是她很少使用社会计量学概念来揭示其内在结构。马林斯(Mullins，1973)则接受另一种策略。他关注到理论社会学中的研究成果，试图揭示专家中存在的子群体。利用关于教育、职位任命以及合作方面的资料，他针对结构功能主义理论、小群体理论、因果理论和大量其他领域的理论构建了社群图[29]。不幸的是，这些专业研究之间的界限本身不是来源于社会计量学分析，因此，马林斯的研究工作很少在理论社会学中给出成分和派系的总体结构的观念。

然而，加特雷尔(Gattrell)的研究是该领域中利用一种严格的社会计量学思路来发现网络结构的少数研究之一。加特雷尔(Gattrell，1984a，b)利用Q-分析技术(参见上述第17条注释)来揭示所研究的群体中的成分结构。这里无须讨论这种复杂程序的细节，因为加特雷尔只是用它来建构一种嵌套的成分模型，并且他的观点可以很容易地翻译为本章的术语[30]。加特雷尔搜索出在1960到1978年间发表的地理学论文，并把它们看成是在空间模型文献方面的关键元素。他把这些论文看成是研究总体，从它们的参考文献和注释中建构了一种引用关系网。例如，只要论文A的作者引用了论文B的作者，就说存在一个从A指向B的引用关系。因此，这些引用数据可用来构建一个二值有向矩阵。由于矩阵的行和列是按照时间排列的，即按照发表论文的时间来排序，因而很容易评价引用模式方面的任何明显的转变。例如，如果作者们只引用近期发表的文献，矩阵中的各个"1"

---

① "无形学院"这个词是英国著名科学家波义耳在1646年左右提出来的，指的是英国皇家学会的前身——由近十名杰出的科学家组成的非正式小群体。300多年以后，美国科学学家普赖斯(Derek Price)在其《小科学、大科学》(1963)一书中首次把非正式的交流群体称为"无形学院"——译注。

项就与对角线接近。"1"项的分布越分散,引用的时间也越分散。任何围绕着对角线的聚类都将支持普赖斯(Price,1965)的假设,即在引文方面存在着"直接效应(immediacy effect)",但是加特雷尔发现,支持该观念的证据很少。

加特雷尔的论文的主要目的是考察引文数据的成分结构,他针对初始阵进行了两种分析:首先,他分析了被引论文(各个行的值)网的结构,其次,他分析了原文(各个列的值)网的结构。如果两篇被引论文都出现在同一篇论文的引文中,则称这两篇论文之间存在一个关系,一个成分就是由一系列论文,即通过一个持续的此类关系链联络在一起的论文构成的[31]。如果两篇被引论文共享的引用者多于一个,它们就以高层次的多元度关联在一起,即可能研究在各个层次多元度基础上的嵌套成分。

加特雷尔发现,在低层次上,49%的论文形成一个单一的大成分。但是在多元度为6的时候,该值降低到7篇。该成分中的这7篇论文形成网络的核。在该群体的中央,有两个被高度引用的论文,即赫德森(Hudson,1969)和彼得森(Petersen,1970)。赫德森被引17次,彼得森被引15次,但是只有8个引文是共享的。这样看来,赫德森和彼得森形成了一个多元度为8、规模为2的成分(根据Cattrell,1984b:447页计算得出)。加特雷尔的结论是:

"总的图景是……一小群被高度引用的论文,其他论文以较低的多元度与这些论文相关联。一个小的论文成分关注出现的等级扩散,其他论文由于被某些引用原创性论文的文献所引用,因而也加入到这个核心之中。"(Cattrell,1984b:448)

如此看来,对成分及其核心的分析允许人们探究科学研究中的权势结构,这种探究说明,科学派系和圈在促进特定的科学观念和研究的增长方面扮演重要角色。对引用模式中嵌套成分的分析突显出那些"明星式的"被引论文,以及它们的星级在多大程度上达成共识。

注 释

1 这是所谓的"小世界"现象的含义之一。参见米尔格拉姆(Milgram,1967),埃里克松(Erickson,1978),林南等(Lin et al., 1978),基尔沃斯和伯纳德(Kilworth and Bernard,1979)的论述。

2 埃弗里特(Everett,1982)指出,他研究"分块"的方法与结构对等性分析中使用的方法相似。事实上,这些程序是完全不同的,而且必须区分开来。下一章将讨论结构对等性。

3 与区分途径的过程一样，其中会存在大量相同的环，这取决于哪一个点作为所命名的环的开始点。在计算一个图中的环数的时候，必须认识到这种双重计算的存在。因此，在后文中，笔者一般将不会对相同的环做出区分，而是只根据一个任意选择的开始点来给它们命名。例如图 6.2 的社群图(1)中只有 3 个不同的环：ABCDA，ABDA 和 BCDB。

4 埃弗里特宣称，他关于“桥”的概念类似于格拉诺维特（Granovetter，1973）的桥概念，后者用该概念来描述那些把关联子群“强烈地”连在一起的一些“弱”关系。

5 应当注意到，一个环成分中的所有点对都通过一个环连在一起，尽管这些环的长度要比用来定义环成分的环长要长。例如，在图 6.2 中，环成分{E，F，G，H，I，J}是根据一个 4-环构建起来的，但是，如点 E 和 F 这样的点只通过一个长度为 6 的环连接到一起。

6 这一点很重要，因为在无向图中的环必须至少连接 3 个点。因此，二方组永远不会是一个环成分。

7 只要这种分析始终会区分出单独一个大成分，埃弗里特建议运用更进一步的一种特定程序来分解图的结构。例如，他认为可以只考虑双向线（即 A 向 B 发出一条线，B 也向 A 发出一条线）。这些是图中最强的线，他指出，只分析这些线就可以确定出该图的最强的结构特征。这一过程在 UCINET 中是通过把有向（不对称的）矩阵当作无向矩阵读取来完成的。该程序对全部非互惠的线都不予考虑。

8 与埃弗里特的“块”观念不同，这种“块”观念建立在图论的早期应用的基础上。

9 当一个“结群”只由一条线连接的两个点组成时，这两个点可能是其他“结群”的成员，此时即称这条线为其他“结”之间的“桥”。这种“桥”观点与埃弗里特的观点不同，也不同于上一章讨论的由施瓦兹及其同事介绍的观点。在社会网络分析中存在的不幸事实是，一些词语经常被矛盾性地用来描述不同的概念。

10 “嵌套成分”这一术语由 GRADAP 小组首次引入到一个特定的程序中，但它有着很广泛的用途。

11 虽然“二值化（dichotomizing）”一词是在 UCINET 操作手册中使用的术语，“切分（slicing）”这一术语却是程序中涉及的最具有描述性的术语。埃弗里特将这一步骤称为“压缩（compression）”。

12 在此，笔者试图以如下方式来推广赛德曼（Seidman，1983）的“核”概念，即这一一般性概念不再用赛德曼的 k 参数来定义。这一点将在下面的讨论中得到详细的说明。

13 注意到这一点很重要：在测量这个程序中某点的度数的时候，只能根据该点与核心部分的其他成员（而不是整个图）之间的关系。

14 我将在下一节中考察凝聚子群本身。赛德曼特别指出他的 k-核概念化与格拉诺维特（Granovetter，1973）关于弱关系的研究之间的相似之处。

15 注意，尽管最小度数（k）只应针对核心中的点来计算，消失之点所占比例却依赖于整个图的总点数。就该图的规模而言，这确保了该向量在一定程度上是标准化的向量。

16 与赛德曼的 k-核观念相对应，我也引入了 m-核这个术语。该术语的优点在于，它也在成分本身和构成成分的核心之间做出明确的区分。

17 我所谓的 m-核是阿特金(Atkin)的 Q-分析方法的基础，该方法是作为图论的对应项发展而来的。在 Q-分析中可以建构一个 Q-接近度(Q-nearness)矩阵，令人迷惑的是，该矩阵中两个点之间的 Q-接近度等于连接这两点的线的多元度减去一。因此，一个其中各点都相互之间以 2 步-接近的成分将对应于一个 $3m$-核(参见 Atkin,1974,1997,1981;Beaumont and Gattrell,1982;Doreian,1980,1981,1983)。Q-分析法的应用在斯科特(Scott,1986)的附录中有讨论。

18 用本章引人的术语来说，这些将是 k-环的 m-核(k-cyclic m-cores)，其中 k 是环长，m 是多元度(multiplicity)。

19 一对关联点只有在平常的意义上才是一个派系，而且派系分析通常只关注规模大于或等于 3 的派系。

20 该公式将点数和线数连接在一起，用来确定在密度计算中所有可能的总线数。该公式是 $n(n-1)/2$，其中 n 代表点数。它界定了所谓的"三角形数"序列。

21 事实上，许多派系检测程序一般不能对有向图进行操作。强、弱派系这两个术语是我个人的创造，是为了与强、弱成分之间的区别相对应而设计出来的。

22 放宽严格意义上的派系成员这个观点的另一个理由是，在 1940 年代和 1950 年代期间，事实证明很难发现能够高效地确定派系成员的运算法则。现在，越来越多的数学知识和计算上的进步已经破除了这个障碍。

23 在 UCINET3.0 版本中有个缺陷，即它的派系测试算法不能运行。这一点在后来的第 5 版中得到了修正。GROUP 程序可读取 UCINET 的文件，并拥有许多派系检测程序。

24 在 UCINET 中有一项用来区分 n-宗派的技术。莫肯(Mokken,1974)也引入了"n-会社(n-club)"的概念，它是一个直径最大为 n 的成分。虽然很少有人尝试进行该方向的研究，我们仍然可以把它看成是对简单成分这种观念的有用的推广。要注意，一个弱成分既是一个 n-宗派也是一个 n-会社，其中 n 值要足够大，以便连接最大数量的点。

25 k-丛(k-plex)概念是沿着与 k-核相似的思路构建的，k-丛也是赛德曼提出来的，这两个概念都以度数为基础。不幸的是，在这两个概念中字母 k 的含义不同。在一个 k-核中，k 是核中点的最小度数值；而在一个 k-丛中，k 值是不需要与一个点连接的点数。

26 事实上，图(2)是一个 2-派系，而不仅仅是 3-派系。

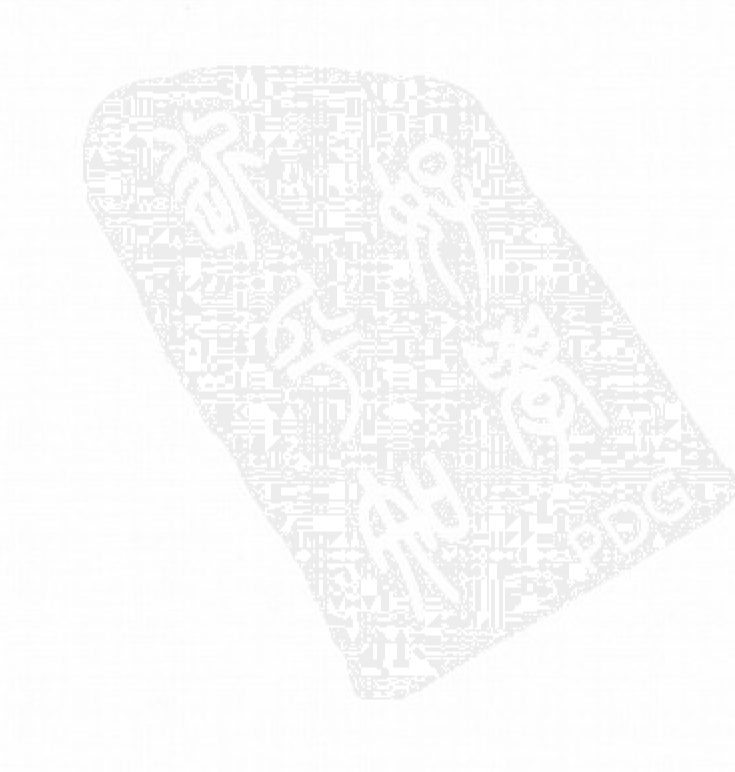
27 赛德曼和福斯特根据莫肯关于限制子图直径的建议提出了另一个概念。他们延伸了 k-丛的观点，并定义了他们所说的"直径为 n 的 k-丛(diameter-n k-plex)"。其定义如下，即它是一组点，其中每个点至少以不大于 n 的距离与至少其他 k-1 个点相连。特别是，"直径为 2 的 k-丛"概念似乎是对派系概念的一个非常有益的扩展。不幸的是，k 测度又是在不同的意义上使用的，因此必须仔细进行对比(参见 Seidman and Foster,1978:69-70)。

28 通常情况下，对环的鉴别可以看作是一种特殊的聚类方法。但在本书中我对聚类作出如下限定，即聚类指的是通过特殊的层次方法发现的子图。环、派系、丛都在点与点之间的相互关联程度基础上对点进行分组，而层次聚类法则考察它们相对于网络其他点来说的整体关系模式。在此意义上，层次聚类法与结构对等性的鉴别高度相关，下一章将讨论结构对等性。

29 马林斯(Mullins,1973)书中第 10 章是对社会网络分析者的研究。它包含一个数据矩阵,但不包括社群图。

30 事实上,加特雷尔(Gattrell)超越了这一点,但我不想在此纠缠这些问题。感兴趣的读者可以参考博蒙特和加特雷尔(Beaumont and Gattrell,1982)的研究并阅读加特雷尔的研究成果。

31 注意到这一点很重要:这并不意味着每一篇论文都与成分中的其他论文直接相连。他不关注派系。数据只显示了一条关系链条的存在。



7 位置、角色和聚类

到目前为止，本书讨论的网络概念主要关注的是行动者相互之间进行直接和间接联系的模式。关注的问题诸如：行动者加入凝聚性社会群体的能力，影响与自己有关联之他人的行动的能力等。然而，我已经在许多点上触及到“位置”分析，而不仅仅是对行动者及其关系进行分析。例如，沃纳和伦特（Warner and Lunt，1942）试图考察各种独特的社会位置的形成，纳德尔（Nadel，1957）也认为，社会角色是社会网络分析的核心要素。最近对该问题的讨论用到的关键概念是“结构对等性（structural equivalence）”。这个概念涉及到由特定类别的行动者维持的社会关系的一般**类型**。尽管两个人可能与完全不同的其他人有直接关系，然而他们与这些人之间关系的类型很可能相同。例如，两位父亲都有不同的孩子集合，但是在某些方面，可以认为他们对待孩子都表现出相同的“父爱”。也就是说，这两个人相互之间是“结构对等”的。他们占据相同的社会位置——即父亲，就对“父亲”进行社会学分析而言，他们的角色是可以互换的。因此，结构对等观念背后的思想就是确定用来界定社会**位置**的行动有哪些一致性。一旦确定了社会位置，就可以探讨各个位置之间的关系网络。

社会位置被一些行动者占据，就他们之间的关系连带来说，这些行动者可以“相互替代”（Burt，1982；Sailer，1978）。在某些重要方面，他们可以互换。尽管社会位置只有在连接能动者之间的特定关系上是明显的，社会位置却不能还原为具体的关系。它们涉及到的关系比较持久，这些关系会随着时间的推移而再生产出来。社会位置之间比较持久的这些关系构成了结构分析的一个独特研究域。

点的结构对等性

初看起来，对结构对等性的分析无非就是对社会角色的分析，但是情况并非如此。两位父亲的例子表明，最清楚的结构对等情形确实是那些产生于人们占据的制度化角色的案例。一个明确的、特定文化角色的占据者是由一些在结构上对等的能动者构成的：他们与相同的他者之间的关系是相同的。当然，这只对于充分制度化的角色来说是真实的。如果人们的行动与标准化的文化期待不一致，他们以各种方式背离他们所臆想或践行的角色，我们就会很少发现行动的一致性。在这些情况下，将不存在与文化上界定的**角色**对应的结构对等性行动者**位置**。毫无疑问，这对于许多文化角色来说确实如此，它们制度化到一致性活动的程度将随着案例的不同而不同。

相反，行动也有可能存在结构的一致性，并且这种行动既不在文化上得到认定，也不在社会角色中得到认同。一些能动者相对于其他能动者来说可能占据一个独特的位置，相对于其他能动者表现出相同的行为方式，尽管这个事实没有被各类参与者认识到。确实，这可能是新角色得以出现的方式之一：新形式的行动得以产生，远在人们开始认识到正在发生之事并赋予其一个名称之前，或多或少得到明确界定的各类行动者之间的关系就开始结晶化。在这个意义上，确定在结构上对等的行动者类别就可能成为区分正在各种突显的角色的一个基础。

因此，重要的是，结构对等性概念可应用于社会位置本身，而不仅仅应用于角色或者原角色（proto-roles）。例如，一个社会阶级恰恰可以根据如下术语进行界定：它是在经济资源的分布方面占据对等位置的一群人，因而相对于其他阶级的成员来说，该阶级拥有在结构上对等的利益和生活机会。

对结构对等的正式讨论开始于洛兰和怀特（White，1971）的影响深远的论文。他们根据角色概念建立了自己的研究，把每个角色的占据者看成是在结构上相互对等的。他们认为，结构对等的行动者在网络中扮演相同的角色，或者与其他位置的占据者保持相同的联系，因此这些行动者之间可以互换。他们将有相同的经验或机会（Burt，1987；Friedkin，1984；Mizruchi，1993）。洛兰和怀特的文章把图论的一些限定性因素描述成一个完整的网络结构模型，并根据代数观念勾勒出另外一种策略。他们认为自

己的研究主要有两个明确的特点，这些特点使得他们的研究区别于社会网络分析。首先，所有点和关系都被同时进行分析，而不是把关注点仅限定在连接这些点的特定线、途径和环上。其次，该研究并不仅局限于对邻接矩阵的分析，而且对初始发生矩阵的行和列进行联合分析。例如，个人及其所在组织可以放在一起分析，而不用单独进行分析。

根据洛兰和怀特的研究，一个网络中各种关系的总体模式必须被转换为结构对等的位置系统，转换方法是把各个点聚集到一个更大的点集之中。相对于在构成这些集合的各个能动者之间存在大量的具体关系而言，这种结构对等系统的内在结构在这些点集之间的关系上表现得更加明显。图 7.1 表明洛兰和怀特的一种观点，即把一个复杂网络“还原”为“块模型”或者“像矩阵”。初始发生矩阵中的点用一种聚类分析的方法进行重排，从而形成了在结构上对等的一系列像矩阵。例如，在图 7.1 中，集合 *M*1 由如下行点构成，即这些行点之间被看成是在结构上对等，但是在结构上不同于构成点集 *M*2 的另一个结构上对等的行点。洛兰和怀特认为，一个网络的最基本特征在各个点集之间的关系中才可以明显看到，并且这些关系的本质可通过像矩阵中各个格值（即块）体现出来。怀特的大多数后续研究的目的都是探讨如何产生这种块模型[1]。

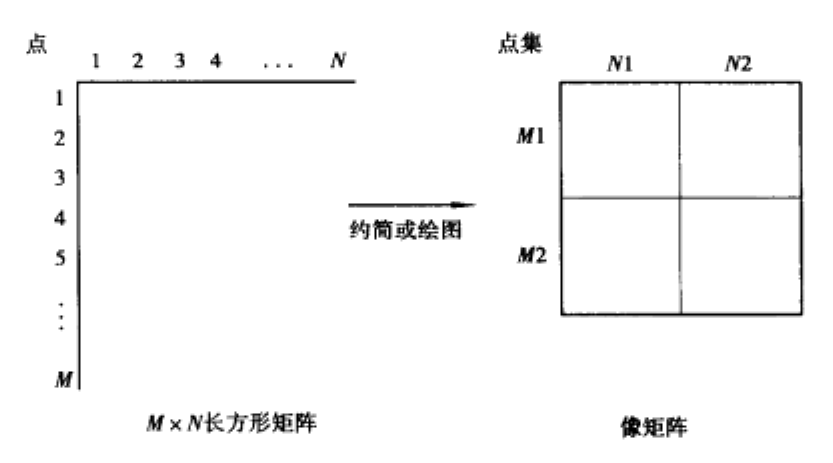


图 7.1　网络及其块模型

就其最强的意义来说，结构对等概念认为，一个集合的各个成员就其与网络中的其他成员之间的关系来说是相互等同的。然而，这种强意义上的完全结构对等很难在现实中见到。因此，大多数结构对等的分析者认为，如果用结构对等性来研究现实中的网络，那么标准应该降低。也就是说，研究的目的不是搜索那些在其社会关系上完全对等的能动者，而是把

足够相似的行动者看成在结构上对等。不管选择何种相似性标准,研究者必须确定一个临界值,在此值之上的行动者被看成是充分相似的,因而可以相互替代。当然,在现实情况下,用来区分对等性的临界水平要由研究者自己确定,其实质效度也需要认真加以评价,尽管如此,结构对等性的这种"模糊"测度仍然非常有用。

结构对等性的研究者们在一些方法上没有达成共识,这主要涉及到所运用的测量"相似性"的特定方法、把点分成小组的聚类方法以及界定群体界限的方法等。在下一部分中,我将简要回顾现有的聚类分析法,并将极细致地阐述两种特定的结构对等性研究。然后,我会再次回到聚类方法的选择上来,还要通过考虑一些替代性的研究方法来考察"相似性"测度。

聚类:聚集和分裂

"聚类(cluster)"和"派系"这两个词经常相互替换使用,在早期的旧城(Old city)和扬基城(Yankee city)研究中,人们对社会计量学中"派系"概念的讨论就是这样。即使是最近一些方法论评论者(参见 Lankford,1974)也没有区分开这两种观念。然而,在上一章中,笔者已经指出,派系这个概念的社会计量学定义是很严格的,从中引申出一系列相关概念。同样,聚类这个概念也需要被明确地界定为一个独立的、非常独特的理念。关于一个聚类的直觉观念是,它对应于图中一个密度相对较高的区域。这种聚类观在关系数据和属性数据中都是适用的,也可以利用诸如图 7.2 那样的散点图加以说明。在一个散点图中,每个个案都根据两个变量来展示,这两个变量构成了图的两个轴。各个点在图中的分布展示了这些点在这些变量上的相似或相异程度。

在图 7.2 中,每个图明显有两个聚类,在(1)中,聚类由彼此之间"相似"的点组成,它们在整个散点图中形成了高密度的区域。虽然这些聚类通过简单视觉审查即可看出来,但是对大型数据库来说,则需要电脑化程序来处理。然而,研究者必须选择一个特定的聚类分析方法。大部分现有的方法都可以区分出图(1)中的聚类,但不是所有的方法都能识别诸如图(2)中的聚类。这个图中的点呈狭长分布,并且每个聚类两端的点之间相距甚远。聚类要根据它们在图中彼此之间的相近程度以及各个聚类之间的远离程度来定义,但在散点图中,不是所有的聚类都由彼此接近的点构成。大多数聚类分析技术都假定聚类是紧密"球形"的,因而很难发现如图

(2)所显示的聚类类型。

显然,聚类之间的边界不能被严格地界定出来。例如,图 7.2 的(3)就展示了两个大的聚类,每个聚类中还包含了一个小聚类。有一种观点认为只有那些小的点集才是聚类。在聚类分析中得到的聚类的构成成分取决于研究者选择的聚类密度标准,也取决于对特定聚类方法所做的假设。

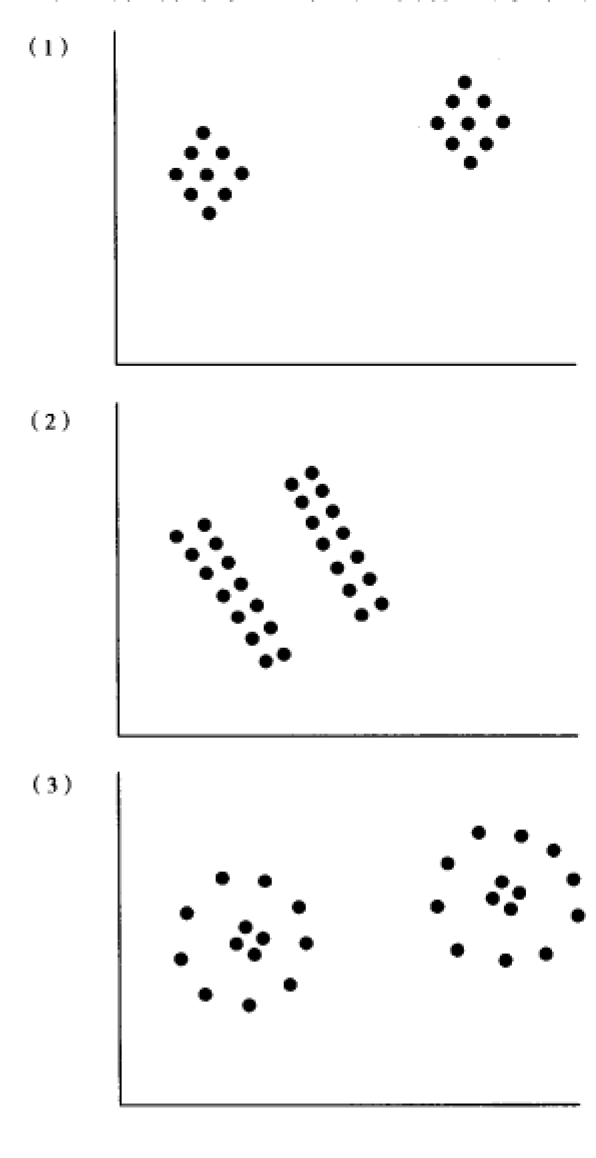


图 7.2 聚类

在决定聚类界限过程中存在的这种主观臆断性表明,可以把各种聚类方法看作是运用嵌套程序的一种变体。聚类有一个等级结构,可以用树形图或树状图来表示,这种图可以展示出在每个相似性层次上存在的各种聚

类。这种观念体现在图7.3中。该图表明,分析的第一步把点C和D连成一个聚类,第二步连接点G和H,第三步把E和F连在一起,第四步连接A和B。如果分析到此结束,那么这四个聚类就确定下来了。如果继续分析到第五步,点E、F、G、H都被定位成一个聚类的成员。具有相似性的点A,B,C和D在第六步中被聚类在一起。最后,在第七步中,所有点都聚合到同一个聚类之中。在研究中发现的聚类的成分和数量都取决于分析所停止的步骤。

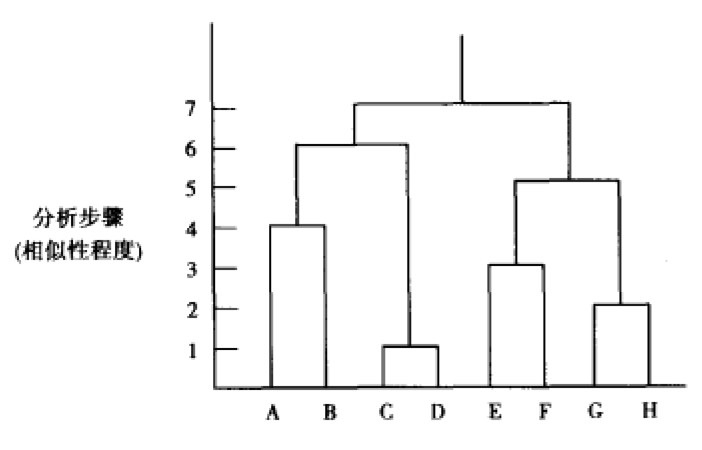


图7.3 树形图

尽管这些观点可以通过我们比较熟悉的属性类型的数据来展示,但是它们也同样适用于关系数据。例如,可以根据点之间的途径距离或者一个社群图的密度来划分聚类。如果采用某种图论意义上的接近性或距离标准,那么一个聚类的各个成员可以是在这些标准上相似的那些个体。

聚类分析方法主要有两类:“聚集性的聚类”和“分裂性的聚类(Bailey,1976;Everitt,1974)”。小聚类嵌入大聚类之中,在这个意义上可以说,每个聚类都具有等级性,但是用来构建聚类等级的原则在这两种情况下各不相同。上文的讨论是根据凝聚性的模型表述的,在这种模型中,单个点逐渐融入到越来越大的点集当中。各个点要根据相互之间的“相似性”或“距离”来比较,并且把最接近或最相似的点分在一组。聚集模型可以采取“单关联(single linkage)”或“完全关联(complete linkage)”的形式(Johnson,1976)。在**单关联**模式中,各个点都和与自己最相近的点聚成一类。例如,在关于连锁董事的研究中,可以根据一些企业共享的董事数把这些企业聚为一类。最初,两个最接近的点组成一个聚类,然后逐渐把联结较远一些的点聚为一类。比如,对于两个聚类来说,如果它们的最相似的成员比集合中其他的任何一对点都相互接近,那么这两个聚类就又可融

合在一起。一种**完全关联**方法遵循同样的一般途径，但是在测量两个聚类相似性程度的时候，该方法不根据它们之间的最近程度，而是根据它们最远的成员[2]。尽管单关联方法趋于把各个点“连接”到现存的聚类之中，然而完全关联法更容易在分析的早期阶段产生新的聚类。因此，单关联法不易于发现紧密的、同质性聚类，而在完全关联法中却可以区分出这种聚类。单关联法强调的是聚类之间的联系，因而能够表明网络中存在的重要分派(Alba，1982：55-6)。

对于两种聚集性的聚类分析方法，分析者需要确定用来区分聚类的相似性程度是多少。在一个关联图中，所有点最终都会融合到单个聚类之中，所以，聚类的数量和规模取决于所选择的截限(cut-off threshold)。与社会网络分析的许多领域一样，可想而知，截限的选择是需要研究者做出有根据的判断的一件事情。

分裂式的或分区式的聚类研究则需要遵循相反的原则。这种研究开始于作为整体的一个图，把该图看成是单一聚类，其中的各个子集合要根据越来越低的相似性水平来分裂。分裂聚类的方法有两种：单一属性(single attribute)法和全体属性(all attribute)法。**单一属性**法开始于如下程序，即把那些占据一种特定指标(indicator)或者特定值的点与不占据这种指标的点分开，因此，根据是否占据这种指标来进行初始聚类。同理，在后续步骤中都遵循这种程序，以便进一步对每个聚类再细分[3]。因此，单一属性程序是由一系列二分法组成的，其目的是产生互斥的点集。另一方面，在**全体属性**法中，初始的分裂及随后的分裂都要以一系列点与图中所有其他点之间的平均相似性为基础。

聚类分析的各种方法似乎比较抽象，含混不清，但是我希望它们的普遍原则是明确的。显然，在一个特定图中区分出来的聚类依赖于所选择的方法以及该方法所运用的相似性测度。如果考察建立在洛兰和怀特(Lorrain and White，1971)著作基础上的一种研究思路的话，那么就可以追溯这句话的含义了。

块模型：CONCOR 和 BURT

根据洛兰和怀特建议的思路进行的第一个可行的探查结构对等性的算法是怀特的两个学生布雷格(Breiger)和施瓦兹给出来的，他们独立地重新发现了麦奎迪(McQuitty，1968)[4]提出来的矩阵聚类法。他们的算法称为

CONCOR[即迭代相关收敛(CONvergence of iterated CORrelations)],这是一个相当复杂麻烦的程序,尽管其一般原则是十分简单直接的。CONCOR算法运算的基础是有关案例和隶属项之间关系的社会计量学意义上的邻接矩阵,可以针对行或列进行分析,或者对行和列同时进行分析。然而,最好只对行进行各步分析,这样才可以很好地理解该算法的一般逻辑。这种分析研究的是各个案例之间的结构对等性。

分析的第一步是计算矩阵中所有的案例对(all pairs of cases)之间的关系,即要测量各个行项值(row entries)之间的相似性。对于每一对案例来说,他们之间的相似性可通过二者之间的皮尔森相关系数值来测量:两个拥有完全相同隶属关系模式的案例之间的相关系数值为+1,而拥有完全不同隶属关系模式的案例之间的相关系数值为-1。第一步分析的结果是一个案例-案例相关系数方阵,这是一种特殊形式的邻接矩阵。第二步要运用一种聚类程序,根据测量出来的相似性值把全部案例分组为结构对等的点集。如果各行是完全相关或者完全无关的,那么分组就比较容易。相关系数矩阵中的所有值都将是+1或-1,因而可利用一种强结构对等性标准把该矩阵分为两个点集。该矩阵因而可归为两个点集,每个点集内部有完备的联系,但二者之间没有任何联系。对于如图2.6所示的数据来说,这样的聚类是可能的。由于在现实生活中通常看不到这种模式的聚类,所以,在确定对等点的"模糊"点集的时候,就必须运用适用于更广范围相关系数值的聚类方法。

通过把行相关系数值转变为一个严密的模式,CONCOR就可以获得一种模糊的聚类。这需要针对每一对案例,计算二者在所构建的相关系数矩阵中的相关系数值之间的相关系数。也就是说,需要计算各个相关系数值之间的相关系数,进而组成一个新的相关系数矩阵。这个过程一直重复下去,即计算相关系数的相关系数的相关系数,一直持续下去。最终会发现,这种重复的相关系数矩阵中所有值都是+1或者-1。这种迭代(重复)的相关系数值表现出一种简单的模式,每行可以分裂为两个聚类,这与上文介绍的应用一个强结构对等标准产生的结果是一样的。每个聚类都由一系列结构对等的案例构成。

利用完全相同的方法,可以把两个聚类细分成更小的构成成分。为了做到这一点,该算法需要回到初始值矩阵,把它分成两个独立的矩阵,每个都对应着已经确定下来的聚类。在第一轮迭代中,把在一个聚类中的各个行成员变成相关系数,再把相关系数作相关系数分析,一直这样持续下去,直到每个聚类中产生了+1和-1这种关系模式。在这个点上,可以又把聚类进行分区,整个程序也再重复一遍。只要研究者愿意,这种聚类的分

区和再分区方式可以一直持续下去，尽管聚类的数量越大，对最终结论的解释就越难[5]。

虽然研究者必须主观地决定何时结束聚类的分区和再分区，每一步骤中 +1 和 -1 值模式的出现确实意味着存在着一种相对无异议的用来确定结构对等的强标准。对每个案例的分区只依赖于在最后矩阵中产生的实际值[6]。不幸的是，这个模式出现的原因仍然很不明确。这意味着，在 CONCOR 算法中执行的是一个未加具体阐述的，也在一定程度上令人费解的聚类原则。出于不很清晰的原因，这种算法本身把原始数据转换成结构对等的各类。因此，由 CONCOR 方法确认的聚类就像根据那种不能产生这种精制模式的程序所形成的聚类一样，都是“模糊”的。

同样，这种聚类分区过程也可以在初始发生矩阵的列中重复进行，这就产生了一个由各种隶属项构成的分组。如果案例代表个体，隶属项指代这些个体所在的组织，那么对组织的分区就是根据这些个体加入组织的模式对这些组织进行的聚类分析。对于初始发生矩阵的各行和各列来说，CONCOR 产生的是一个等级式的分区，分成各个结构对等的聚类——即分区成为离散的、互斥的、完全的各类（knoke and kuklinski，1982：73）

用这些方法确认的聚类可构建成如图 7.1 所示的重排过的像矩阵。可以为个案邻接矩阵建立一个像矩阵方阵，也可以为隶属项邻接矩阵建立一个像矩阵方阵。在像矩阵中，每个单元都称为“块（blocks）”，它们包含各个集合之间的关系密度值。如果所有的密度值不是 1 就是 0 的话，那么关系模式就是清晰明了的。“0-块”（zero-blocks）（密度值为 0 的一些单元）代表网络的“洞（holes）”，即完全不存在关系；密度为 1 的单元的分布表示了网络的基本结构。这种密度模式很少发生在真实数据之中，因此，一种块模型必须把实际的密度值范围转换成高、低这两类，以便作为对 1-块（1-blocks）或 0-块（zero-blocks）的近似。在像矩阵中，高密度值（即大于某个特定截限值的值）用 1 表示，低值用 0 表示。在界定高密度块的时候，最常用的方法是把整个矩阵的平均密度值当作衡量标准：大于或者等于这个平均值者就是“高”值，而低于这个平均值的就是“低”值。但是，与网络分析中的许多程序一样，这个程序也涉及到研究者自己的谨慎选择，并且这个选择应该建立在理论或者经验考量的基础上。我们不能仅用某些纯粹形式的、数学原则来判定。弗里德金（Friedkin，1998：8）也反对把密度作为衡量块结构的唯一标准。

完全相同的程序也可用于对行和列同时进行分析，从而产生一个块模型。CONCOR 会形成一个行聚类和一个列聚类，然后把这两个聚类结合到关于初始发生矩阵的一个单一像矩阵中。

一旦生成一个块模型（这是其各项中只有1和0的一个像图），研究者就必须试图解释它。对来自于长方阵发生阵的块模型的解释是相当困难的，而作为块模型的开创者，布雷格及其合作者也没有对这些模型作任何详细的分析。在对邻接阵的最早分析中，布雷格、布尔曼和阿拉比（Breiger, Boorman and Arabie, 1975）重新分析了戴维斯和他的同事于1941年收集到的美国东南部地区（Deep South）的数据，该数据基于18位女士参加14起社会事件[7]。为了分析这些数据，他们对行和列进行了独自的计算，然后又把它们结合到如图7.4所示的一个块模型中。

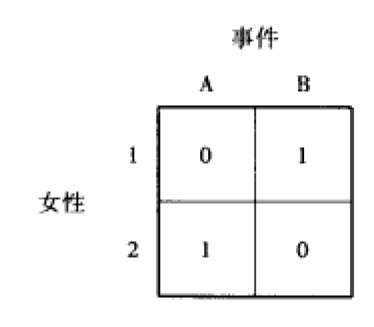


图7.4　简单的块模型

由图可见，聚类1中的女性倾向于参与聚类B的事件，聚类2中的女性倾向于参与聚类A中的事件。这两个女性聚类很接近于霍曼斯（Homans, 1951）在分析初始数据之时确认出来的两个派系，但布雷格等学者并没有超越这一观察视角。虽然他们讨论过聚类的构成，但他们并未注意到像矩阵中块密度的模式。在同一篇文章中，他们也对莱文（Levine, 1972）的有关银行和公司之间关系的长方形矩阵作了重新分析，但是他们仍然与莱文一样，只是简单地对行和列进行单独的分析。

块模型的发明者不能对发生矩阵进行详细分析，这表明，在洛兰和怀特所期待的对行和列进行同时分析方面存在着一种根本性的困境。如果一个长方形的像矩阵很简单，它就可以给出一个网络的初始的、纲要式的概要，但是，更详细的分析只能通过分别分析行和列来实现。如此看来，对一个发生阵来说，必须分别对构成它的邻接矩阵进行块模型的构建，这才能分析该矩阵。在这些块模型中，含有1值的对角线单元对应着上一章讨论过的派系或者社会圈。其他单元表达的是构成图[8]的各种派系和聚类之间是否存在联系。

布雷格已经指出如何在社会网络分析的一些核心领域中利用CONCOR方法。他运用来自英国（Glass, 1945）和美国（Blau and Duncan, 1967；Featherman and Hauser, 1978）的社会流动数据，建构了一个阶层结构模型，其中阶层被定义为一个职位流动率矩阵中区分出来的职位集合（Breiger, 1981, 1982）。他把这看作是对韦伯（Weber, 1920—1921）的如下命题的一个扩展，即"只有当阶层内部个体的流动机会以诸如能够产生共同的社会互换关系的形式聚类在一起的时候，社会阶层结构才能存在（Breiger, 1982:18；同样参见Scott, 1996:第2章）。他建议，CONCOR可用来界定阶层的范围。布雷格用成年男性的代际流动矩阵，其中一个是美国

的 17×17 有向的职业类型矩阵,另外一个是英国的 8×8 有向矩阵。在每个矩阵中,单元格中包含了从一类职业向另一类职业转变的数量。行值表示"起点",而列表示"终点"。布雷格(Breiger,1981)认为,对于美国数据来说,在 1962—1973 年间八个阶层存在着稳定的结构,而对于英国数据,他的结论是,早期的数据(1949 年左右)反映了三阶层结构。在英国,核心阶层(central class)边界把体力劳动和非体力劳动分开,把工薪阶层的中产阶级与低层次的文书与行政工作阶层分开。

到目前为止,最早利用块模型进行分析的矩阵是有向数据的邻接矩阵——例如,行值表示"发出"的关系,列值表示"收到"的关系。用来解释这种数据的有用手段便是构建聚类之间关系的箭头图(arrow diagrams)。图 7.5 可以说明这点,该图展示的是假定的权力关系[9]数据。在这些矩阵中,权力关系是从行指向列的。例如,初始阵中的行值展示的是行使权力之人。相反,列值展示的是在一个权力关系中,一个特定的能动者行使权力的对象。在块模型中,要根据使用权力和接受权力来分类,"像矩阵"中的"1"和"0"项即表达了聚类之间的权力关系密度值。

在图 7.5 的模型(1)中,聚类 1 的成员相互之间行使权力,同时也对聚类 2 的成员运用权力。这一点可以在相关块中的"1"值得到展示。然而,毫无疑问,聚类 2 中的成员没有权力可以行使,相对于聚类 1 的权力来说完全处于屈从地位。这个结构汇总在相应的箭头图中。另一方面,在模型(2)中有两个独立的、自我控制的类别(聚类 1 和聚类 3),这些聚类中的成员共同对聚类 2 的成员运用权力。最后,在模型(3)中,聚类 1 控制着聚类 2 和聚类 3,但是聚类 1 的成员之间很少相互行使权力——聚类 1 的每个成员都是相对自主的能动者。

除了一些最简单的案例之外,无向矩阵是很难解释的,因为没有任何方向的关系意味着不可能建构箭头图来表示它们的结构。有关这方面的分析成果很少见,如果要展示块模型的价值,那么应该在各类真实的、复杂的数据库中运用块模型,这才是重要的[10]。

笔者已经指出,CONCOR 算法的一个基本问题是我们不能确切地知道产生解的原因是什么。相关系数矩阵经过迭代之后汇聚于 1 与 0 分布的数学原因仍然不明确,因此,结果的效度难以评定。这样的批评似乎是毁灭性的,但事实是它确实起作用了,并且似乎产生了可信的小社会网络模型,这在一定程度上抵消了这个批评。然而,还有个困难更进一步限制了它的应用性:CONCOR 只有在一个图的成分和子群体的**内部**才能确定结构

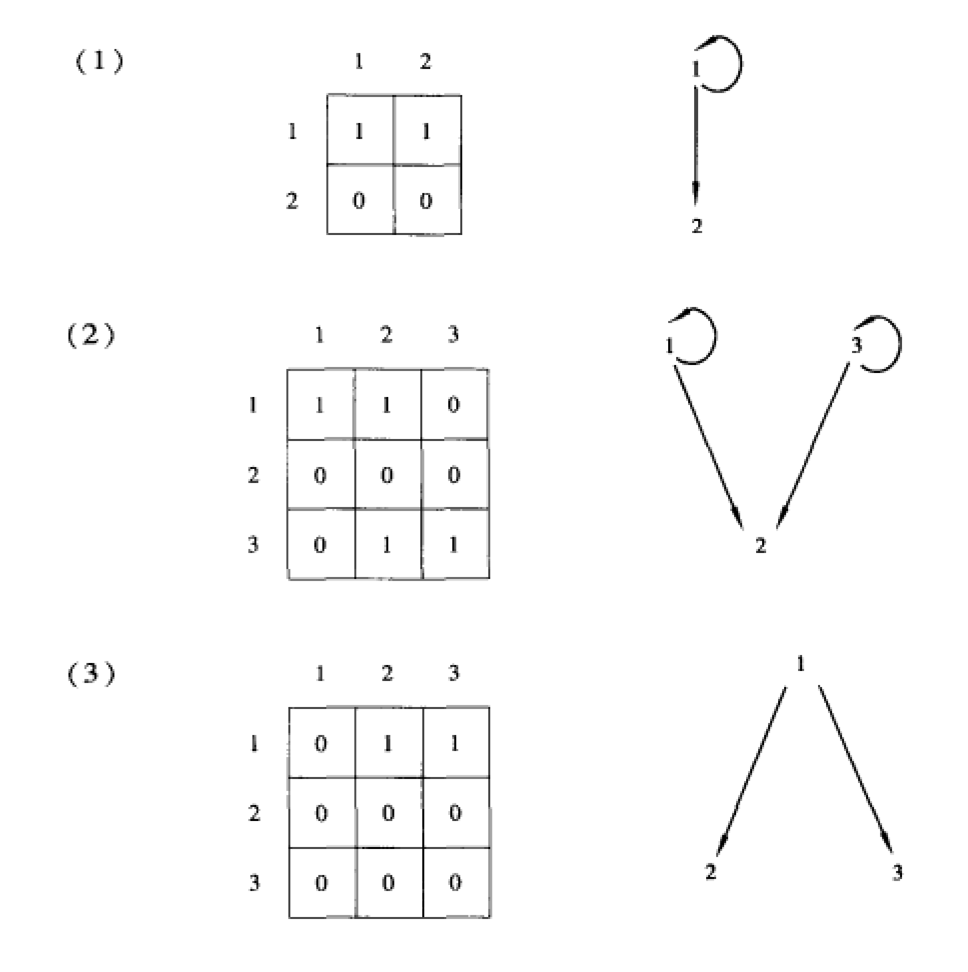


图 7.5　等级块模型

对等的位置。例如,如果网络中的权力关系被分割到每个成分之中,那么 CONCOR 就无法把那些在各个独立的成分中处于支配地位之人再组成单一的居于"主导性"的能动者聚类。作为支配地位的占据者,他们的对等性在 CONCOR 中被掩盖了,即被他们在社会计量学意义上分成的各个独立成分所掩盖。同样,当一个成分从内部被分为一些相对独立的派系和圈的时候,CONCOR 就会有一种只能区分出每个子群体内部的支配成员的倾向。

上述观点可通过如图 7.6 所示的网络加以展示,在该网络中,A、B、C、D 都是占据支配地位的能动者,因而在结构上是对等的。E、F、G、H 是具有结构对等性的被支配者。如果把该网络充分聚类成为结构对等的位置,那就应该区分出两个聚类{ABCD}和{EFGH}。然而,如果能动者被分成如图所示的两个不同的社会圈,那么 CONCOR 就会区分出四个聚类:{AB},

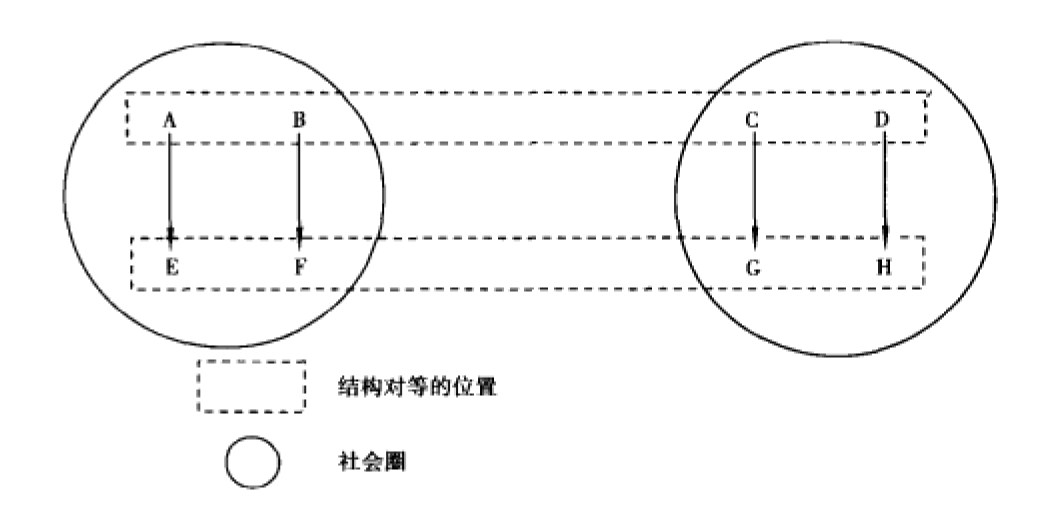


图 7.6 结构对等和社会圈

{CD},{EF}和{GH}。由此得到的结论是,CONCOR 算法把结构对等性与传统的社会计量测度结合起来,却不能对结构对等性进行彻底的分析。

CONCOR 程序的这些局限性促使罗纳德·伯特(Ronald Burt)提出另一种结构对等性研究思路,其目的是避免 CONCOR 方法对一些不确定的数学程序的依赖。无论从所利用的相似性测度,还是从聚类方法角度讲,伯特的方法都不同于 CONCOR。在考察完 CONCOR 的一些细节后,就很容易理解伯特的程序是如何运作的,以及它在哪些方面对 CONCOR 进行了改进。伯特的程序在计算机程序 STRUCTURE 运行,为了把它与运行的其他程序相区别,我称之为"BURT"[11]。

BURT 所用的相似性测度根据的是两点之间的途径距离。CONCOR 只根据直接联系来测量相似性,BURT 还考虑到距离为 2 或更多的途径的间接联系,以便计算出所有点对之间的最小途径距离值。伯特式的途径距离测度也假定关系的强度会随着途径距离的增加而降低,也随着途径对于能动者的总关系模式的重要性的降低而下降。这个测度建立在如下假设之上,即与大量他者有交往的能动者较少注意到距离他们较远的一些人[12]。因此,伯特中所运用的相似性测度是一种加权的距离测度[13]。

从严格意义上讲,结构对等的能动者之间的距离是 0。它们完全相似,并且可相互替代。伯特认识到,这种强标准不能用到大部分现实的数据中,所以他坚持确定弱对等性,这就需要利用一个临界距离值,低于此值的点即被认为是结构对等的(Burt,1980:101 页及以后)。尽管 CONCOR 的任意性来自于一种不明确的数学程序,BURT 的任意性却有一个优点,即它

建立在专业研究者的有见地的判断之上。BURT 针对距离矩阵运行一种等级聚类，需要利用约翰逊（Johnson，1976）的聚集性的单一关联法（aggregative single linkage method），研究者并且根据所选择的临界距离值，对发现的聚类进行取舍。

一旦利用 BURT 程序产生了一个聚类，接下来就需要构建块模型。如果利用整个网络的密度值作为临界值，用 1 或 0 来代替像矩阵中的各个密度值，那么得到的块模型就可以像在 CONCOR 分析中一样进行分析。用来表达各个聚类之间关系密度的像矩阵是一个简化了的图，是对各个能动者之间的具体关系模式的一个同形还原（homomorphic reduction），伯特称之为一个"社会拓扑（social topology）"。

伯特认为，从一个严格的强结构对等测度中脱离出来，这意味着任何分析都应该仅仅被看成是一个假定的模型。他认为，如果没有某种统计显著性检验，研究者就可以自由地选择任何截限值，从而使结果接近于他们的预想。显著性检验有助于在对块模型的评价中引入一定程度的无偏性和客观性。伯特推荐的检验涉及到对每个聚类进行检验，以便测量聚类的每个成员与其他成员之间在多大程度上接近。他指出，最好的解决方法便是对这种相关测度进行优化处理[14]。

走向规则结构对等性

在确认结构对等性的时候，CONCOR 和 BURT 可能是得到最广泛应用的两种程序，但除此之外还有其他可选项。尽管其中的某些程序已经相对易于得到，它们却很少被应用于真实数据之中，它们的长远价值还有待评价。

CONCOR 只考虑到长度为 1 的途径距离，BURT 则要考虑到所有的关系，不管其长度是多少，而塞勒（Sailer，1978）提出一个程序，在该程序中，研究者可自己选择途径的长度。一旦选定了途径的距离，点对之间关系的相似性程度就要根据选定的长度来计算。然后，赛勒的测度——他简单地称之为"可替换性（substitutability）"——建立在各个点的邻点基础上。两点之间的相似性程度要根据比例值（而不是绝对值）进行测量，即在一个特定距离长度的途径上它们共享的关联点数要根据每个点的邻点进行标准化处理。也就是说，诸多接触点之间的重叠程度要根据它们共享的关联数与每个点在该距离上的总关联数之比来衡量。因此，每个点与其他点之间

都有一个标准化的相似性测度。如果关系完全重叠,标准值就是1,如果完全不重叠,值[15]就为0。如在CONCOR程序中一样,赛勒把这仅仅看作是迭代过程的第一步。相似性矩阵被看成是对点的"可替代性"的最初估计,而这种方法在每个新的估计值上的持续重复会导致一个解,即其中所有值不是"1"就是"0"。这样就产生了需要加以分析的块模型。然而,赛勒的程序也不能克服CONCOR与BURT方法中的原则性缺陷,即它们不能充分地处理被分成各个成分的网络,也不能处理如图7.6所示的那类紧致子群(tight subgroups)(也可参见Carringten and Heil,1981;Wu,1984)。

在克服这种社会计量学意义上的缺陷的时候,一个有趣的尝试是REGE,它是一种致力于探查"规则(regular)"结构对等性的算法。它的定义是,如果一个网络的全部子群都是规则的,这样的对等性就是规则对等性(Reitz and White,1989;White and Reitz,1983;Winship and Mandel,1984)。如果从一个社会系统中的能动者所扮演的角色或功能的角度来看,规则对等性这个概念和"可替代性"观点就接近。一个图中的某点如果与图中其他点之间有**等同**的关联(links),CONCOR和BURT就认为这样的点在结构上对等,而如果一些点与那些本身具有结构对等性的点有**相似**的关系,REGE就认为这些点是对等的。对于两个点和另外一个点集来说,如果两点中的一点与上述点集之间的关系与另外一点与该点集之间的关系类似,则说这两个点相对于该点集来说是对等的。每个点与其对应点与同一个点集的关系完全相同,尽管这种关系不一定是与同一个点或与多个点之间的关系。这个观点可以用一个明显的事实来描述,即所有的父亲都和孩子有关系,但他们并不都是和**同样的**孩子有关系。因此,怀特和赖茨(Reitz)认为,由REGE产生的块模型是同形的还原(homomorphic reduction),但不一定是对相应图的自同构还原。

REGE的运作方式最好通过一个有向矩阵案例来解释,尽管很难理解该程序的具体细节。该算法运用一种分区法,该方法关注的是有向关系,也关注与每一对点相连的其他邻接点。该程序首先要估计全部关联点对之间的对等性取值。这些估计值最初全部设定为1,然后在每一轮计算中不断修正,这涉及到根据每一对点的最小点入度和点出度值来计算各个修正的对等性取值。因此,在每一轮的最后,各个点对之间就产生一个由估计对等值构成的新矩阵。在理想情况下,该程序一直持续下去,直到修正的对等估计值不再改变;也就是说,各个计算结果不会再比估计值更精确。在实际情况下,当更进一步计算的结果与估计值差别不大的时候,研究者就可以停止计算了。有人建议,经过三轮计算之后,UCINET中挂着的REGE程序就会产生最佳的估计值(也可参见Borgatti and Everett,1989;

Borgatti et al. , 1989)。

这种方法只能用于有向数据,尽管为了分析无向数据多瑞安(Doreian, 1987)已经对它作了修正。如多瑞安所言,对一个对称矩阵来说,最初的估计值不会受算法影响:算法无非是把所有关联点都看成是规则对等的。这些矩阵只有被分成两个不对称矩阵时才可用 REGE 进行联合分析。多瑞安建议利用中心度的值进行区分,尽管埃弗里特和博加提(Borgatti,1990)也指出任何一个图论属性都是可用的。例如,如果用中心度,那么一个矩阵就由那些从中心点指向边缘点的关系组成,而其他矩阵则由那些从边缘点指向中心点的关系组成。

尽管 REGE 有一些缺陷,但它却是第一个结构对等性分析程序,它提供了对洛兰和怀特(Lorrain and White,1971)描述的规则对等性的真正近似。然而,关于数据的实质性的假设被复杂的数学程序所掩盖,变得模糊不清了,因而对于一个不精通数学的人来说,很难确定这些假设是否有效,是否现实。至于 CONCOR,它**确实**如期望的那样适用于小范围的数据分析,这个事实有力地维护了它的声名,但是研究者必须认识到自己只是在某种程度上信任该程序罢了。

可以设想,诸如纳德尔这样的学者的愿望是建构一个社会学的分析框架,在该框架中,对地位的分析会补充比较传统的社会计量学意义上的派系和成分分析。在导致 REGE 的一些结构对等性分析途径都避开了图论,因而距离这种社会计量学关切只有一步之遥。另一方面,"图角色分析"(graph role analysis)的途径试图运用在图论中测量的各个点的结构位置作为测量结构相似性的基础(Zegers and ten Berghe,1985)。该程序运用局部依赖性(local dependency)或者捷径矩阵(geodesic matrix)来计算每一对行动者之间的相关系数[16]。评价结构对等性的根据是这些测度对于每个点来说在多大程度上相似。例如,从某些重要的方面来说,有着同样高的中间中心度指数(betweenness scores)的一对点就可被看作是结构对等的。如果只要两个点处于**相同的**很多点之间,就认为二者在结构上对等,那么这里明显存在一些问题,为了避免出现这种明显的问题,该算法可以计算这些点是否处于那些在中间中心度取值上具有相似性的两点之间。这个程序的一个特别有趣的特征是,它开始在相对易于理解的图论概念和不易于理解的结构对等性测度之间建立一个桥梁。与 CONCOR 一样,该程序的目的不在于把这些研究合并在一起,而在于从理论上思考二者之间的互依性,并把它清晰地表达出来[17]。

连锁与参与

伯特对商业领域中的连锁董事(interlocking directionship)问题一直感兴趣,但是他的研究没有利用基于派系的传统研究方法。他在关于该问题的最早研究(Burt,1979)中就热衷于揭示在公司系统中的利益与企业的结构地位之间的关系,他提出的结构对等性观念面对的恰恰是结构地位问题。

伯特开始于如下假设,即可以把很多连锁看作是"增选机制(cooptive mechanisms)",通过这种机制,某些企业会把其他企业中的那些可能会威胁到本企业持续运行的人吸引过来。因此,产生市场不确定性的那些供应商就是向它们提供商品和资本的那些企业"增选连锁(cooptive interlocks)"的对象。因此,金融机构在企业联盟中是特别重要的:"作为一种通用的资本,资金的使用使得金融机构的行动成为不确定性的一个重要来源,因此,可以期待公司建立与金融机构之间的增选连锁关系,这样就可以保证在需要的时候可获得资金。"(Burt,1979:416)

在早期对"位置"概念的讨论基础上(Burt,1976,1997a,b),伯特把运行于各个经济领域的公司看成是相互结构对等的——各个经济领域是由社会拓扑结构(social topology)中的各种地位构成的。利用1967年的美国各个部门层次的输入-输出数据,伯特试图展示,在哪些部门之间的交换领域存在的不确定性程度有多大,从而使得增选联盟成为一个理性策略。也就是说,他的兴趣点在于弄明白,限定性的经济交换结构是否反映在与之对应的联盟结构之中。两个部门之间存在的"张力"这个概念要根据竞争压力来操作化:对于一些企业来说,它们既可以与垄断性的部门进行交换,也可与竞争性的部门进行交换,相对来说,这些企业的交换更多地受到前者的控制。市场中的控制降低了企业的结构独立性,而联盟降低了这种控制的影响,因而转换了企业运行的经济环境。伯特坚持认为,"两个网络中的结构是一种共生现象(原文如此):市场结构使连锁结构模式化,而连锁结构又把市场结构再模式化(Burt,1979:433;同时参见 Burt et al., 1980;Burt,1982:第4章和第8章)"。

伯特的数据由两个类似的有向邻接矩阵构成,其行与列对应着各个经济部门。其中的一个矩阵包含了两个部门之间的经济交换信息,而另一个则展示了它们之间的连锁模式。伯特并没有汇报对这些数据进行块模型

分析的结论,但他的结论是,这两个网络确实十分相似,并且很可能区分出一个"董事连带市场(directorate tie market)"——这种连锁结构为商业交易的规范化提供了一种"非市场(non-market)环境"(Burt,1983b,c)[18]。

CONCOR 程序的优势和缺陷都在我主持的公司合作持股研究中明显存在(Scott,1986)。笔者选择 1976 年的 250 个英国最大的公司作为研究对象,并根据他们的股份登记表确认出最大股东。这就构建了一个关于各个公司之间交互持股的 250 × 250 发生阵。在这个矩阵中,行值表示控股公司,列值表示持股关系的对象:控股关系是从行指向列的。研究发现,只有 69 家公司在其他大公司中持有具有控制性的股份,而只有 140 家是这 69 家公司的持股对象。因此,有效的数据库是一个 69 × 140 矩阵。中心性分析表明,宝德信保险公司(Prudential Assurance)是最核心的持股参与者,在 140 个目标企业中,它持有其中 88 个公司的股份。同样,研究发现,药物零售商布思公司(Boots)是目标中最"靠得住的(blue chip)"——在其最大的 20 个股东中,有 18 个就职于最领先的 69 个公司中。

这种分析的主要目标是运用 CONCOR 算法来解释公司网络的一些总特征。有控制能力的公司被看成是经济生活中的主要能动者,该研究的目的是弄清楚它们是形成了一个统一的群体,还是分裂成竞争性的与团结性的联盟。成分分析表明,这些企业之间很少组成联盟,结论是网络并没有分裂成各自独立的公司集群。然而,CONCOR 揭示了在网络中存在许多结构位置,与之伴随也发现了等级关系的存在。对行与列进行联合分析表明存在着五群企业,如图 7.7 的箭头图所示[19]。

在图 7.7 中,箭头表示两个不同集群企业之间的持股关系方向,这些不同集群企业是由网络中的各种位置构成的。集群 1,集群 2 和集群 4 合在一起,构成了经济领域中的支配控制者(hegemonic controllers)。而集群 1 在这个分组中占支配地位,集群 1 包含 20 个企业、一些大的公共部分企业和商业银行,它的成员是其他三个集合的主要持股参与者。它占据着公司等级的最高位置,因为其成员由富裕的家族(families)和国家(而不是由其他公司)控制着。集群 2 包括 11 个企业,这些企业相互拥有资本(这一点用自我指向的箭头表示),它们同时是控制集群 3,4,5 的重要参与者。集群 4 与集群 2(它由交换银行、保险公司和大型私营企业组成)很相似,但是其成员较少卷入到对集合 3 的公司和财团的联合控制之中。这是二者的区别所在。集群 5 包含了 91 家企业,由一些次要集团组成,这些企业实际上不控制其他企业。由于这些英国企业网络没有内在的分散性,CONCOR 在揭示这些企业所占据的结构对等性位置方面是很有效的。但是,利用来自日本的类似数据就发现 CONCOR 的用处不太大。日本的经济

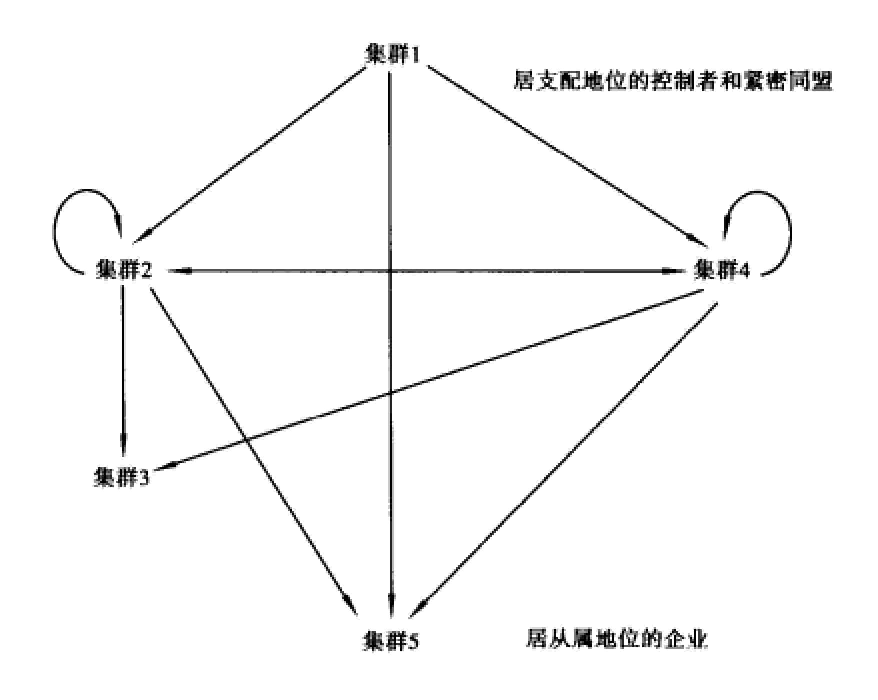


图 7.7　1976 年英国金融统治的结构

被断然地分割成各个独立的部分,每个部分都像十分有凝聚力的商业组织那样运作。这些就是人们熟悉的日本商业体系中的"商业财团(kigyoshudan[①])"(Scott,1991,1987)。尽管存在一些具有结构对等性的主导产业和次属产业,CONCOR 还是根据商业集团对这些企业进行区分(Scott,1986:第 186 页及以后)。利用 CONCOR 并没有发现单个集合的结构对等的统治者。在该网络中发现有 7 个集合,其中 3 个对应着著名的住友(Sumitomo)、三井物产(Mitsui)和三菱(Mitsubishi)这三个商业财团(集团)。每个集合都是有等级划分的,可划分为支配性的企业和被支配的企业。因此,日本的经济看起来很像图 7.6 所示的那样,其结构对等的各个位置被那些代表主要商业集群的社会圈从中间断开。

① Kigyoshudan 是根据日文中的"商业财(集)团(business groups)"一词的日语发音译出来的——译注。

注 释

1 正如在社会网络分析中的许多领域一样，这里也有混淆不清的术语问题。继洛兰和怀特之后的研究者用“块”一词既表示点的集合或聚类，也表示像矩阵中的各项。第六章已经指出，图论专家用“块”一词来指称大量完全不同或无关的概念，这个事实也带来进一步的混乱。在该章规避了“块”一词后，笔者在本章中作了一些限定，即用“块”来描述像矩阵（image matrix）的各项，而将点的集合称为“聚类”，其原因将在下一节中指出。

2 也有其他形式的聚合性的聚类分析，但这些都是最常见的。一个折中的办法是在UCLUS算法中使用平均距离法。在选择相似性测度的时候，它既不挑选两个聚类之间的最短距离，也不选最大距离，而是选用平均距离。安德伯格（Anderberg，1973）提供了一种中间性的“平均关联”方法。也可参见艾伦（Allen，1980）。

3 有些作者建议用卡方值（χ^2）来选择每一步的标准。

4 关于此方法的最精确的解释可参见布尔曼和怀特（Boorman and White，1976），布雷格等学者（Breiger et al.，1975）以及怀特等学者（White et al.，1976）的研究。最初的方法在麦奎迪（McQuitty，1968），麦奎迪和克拉克（McQuitty and Clark，1968）的文章中有讨论。

5 也有人认为，将矩阵拆分为包含少于3个点的集合是无用的，因为在多数矩阵中，规模为2的一组只是关于结构对等性的一种意义不大的情况。

6 事实上，有学者发现在所有值汇聚在+1或-1之间以前，该进程就可以停止。有人指出，虽然在设置高于+0.9和-0.9的汇聚标准上耗费了额外的计算时间，但是所取得的收效却甚微。在UCINET中运行的CONCOR允许研究者自己选择汇聚的标准。然而，一旦做出一项决定，便会理所当然将一种随机因素引入到程序之中。

7 布雷格等学者（Breiger et al.，1975）的分析与戴维斯等人（Davis et al.，1941）的数据无关，而是对霍曼斯（Homans，1951）给出的初始数据的再分析。

8 当一个块的真实密度为1.0时，该块就是一个1-派系。然而当其密度小于1时，块将不会形成一个1-派系，甚至不会形成一个n-派系。这并不意外，因为CONCOR是专为代替派系检测方法而引入的。如果CONCOR仅仅识别派系，那么它的复杂程序就没有什么意义了。

9 布雷格（Breiger，1979）的书中有关于更复杂但是类似的真正数据集合的讨论。

10 关于CONCOR和块模型的更多文献可在莱特和马林斯（Light and Mullins，1979）、阿拉比等（Arabie et al.，1978）以及施瓦兹（Schwartz，1977）中找到。CONCOR程序在伯纳西茨和麦康纳基（Bonacich and McConaghy，1979）中得到了推广。还可参见卡林顿等（Carrington et al.，1980）。

11 这一讨论依赖于伯特（Burt，1980，1982）、伯特和肖特（Burt and Schott，1990）的观点。怀特的最新著作走向与伯特类似的概念导向。参见怀特（White，1992a，b）。

12 伯特称之为"频率衰减(frequency decay)"假设。STRUCTURE 程序允许在研究者作选择的时候使用另外一些假设。

13 伯特计算了"欧氏"距离,我将在下一章中讨论它。

14 由伯特推荐使用的实际检验包括对每一个聚类的"协方差矩阵"运行主成分分析。一个点在第一个主成分上的负载值被看成是该点与聚类中其他成员之间的关系测度。除非理解了主成分分析法,否则这种检验就似乎有些难懂。这些将在下一章中讨论,读者可以在阅读下一章的相关部分后再回来阅读这一节。伯特的方法在福斯特和罗姆尼(Faust and Romney,1985)中受到了批判。

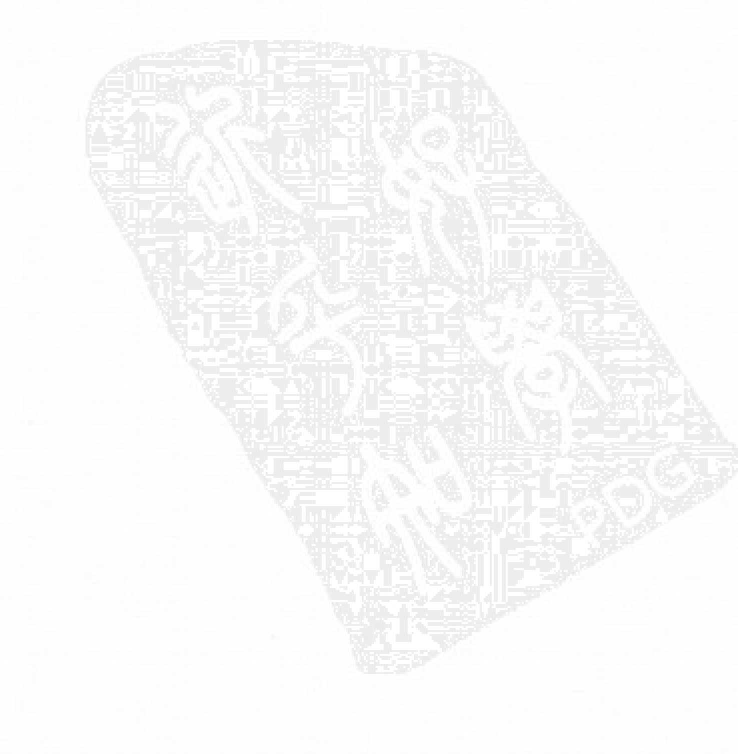
15 由于每个点都有不同数量的联系点,相似性矩阵因此是不对称的。

16 计算出来的相关系数不是一个标准的皮尔森相关系数,而是一个合并了伯特方法中使用的某些欧几里德假设的系数。

17 至于尝试具体探讨结构对等性和传统图论所关注的问题之间的关系,参见埃弗里特等(Everett et al. , 1990)。

18 伯特已将结构对等性与多维量表法(DMS)结合在一起,在伯特(Burt,1988)、伯特和卡尔顿(Burt and Carlton,1989)中生成了美国市场拓扑结构。下一章将讨论多维量表。

19 一个较小的、与网络中其他部分之间联系非常弱的第六集合被排除在外。



维度和展示 8

社群图——网络图——是最早对社会网络进行形式化分析的工具之一,并且社群图法已经成为发展并展示社会网络分析概念的一种重要技术。正是出于这种目的,本书大量地使用了社群图法。例如,中心度可以用某种社群图来展示,在该图中,一系列发散的"辐条"把一个中心点与一些边缘点连在一起。但是,作为一种代表并展示关系数据的方法,传统的社群图法有一定的局限性。主要局限性在于,在一张纸上画一个大图是很难的,这一点限制了社群图的应用。如果图中之点超过 10 个或者 20 个,即使对于密度相对较小的网络来说,交叉的链条数也会导致不可解释的线条集合。

为了克服这个局限性,研究者试图利用某类图形表达方法来补充这些数学测度,从而对社群图观念进行了各种特定的推广。其中一种常见的技术是沿着一个圆的圆周建构社群图,从而使得各条线之间的关系模式更加清晰可见(Grieco,1987:30)。图 8.1 就是此类方法之一例,该例取自斯科特和休斯(Scott and Hughes,1980)。圆的使用仅仅是作为一种用来组织数据的、具有任意性的可视性框架,各个点在圆上的排列顺序仅仅取决于要保证连接这些点的线之间的重叠性最小。研究者参与到一种试错,即指派和再指派的过程当中,直到取得一种在美学上令人满意的解决方案来[1]。

这种图常常可以使一系列关系的结构更清晰,但它们仍然是具有相当随机性的排列,并且没有体现出什么特殊的数学特点。各个点被随机地安排在一些位置之上,画出的各条线的长度反映了这种随机性的安排。对数据的解释相对于展示出来的特定可视性构型(visual configuration)来说在多大程度上具有敏感性,这一点在麦格拉斯等学者的论文(McGrath et al.,1997)中已经指出。在一项实验研究中,人们仅仅通过利用不同的空间安

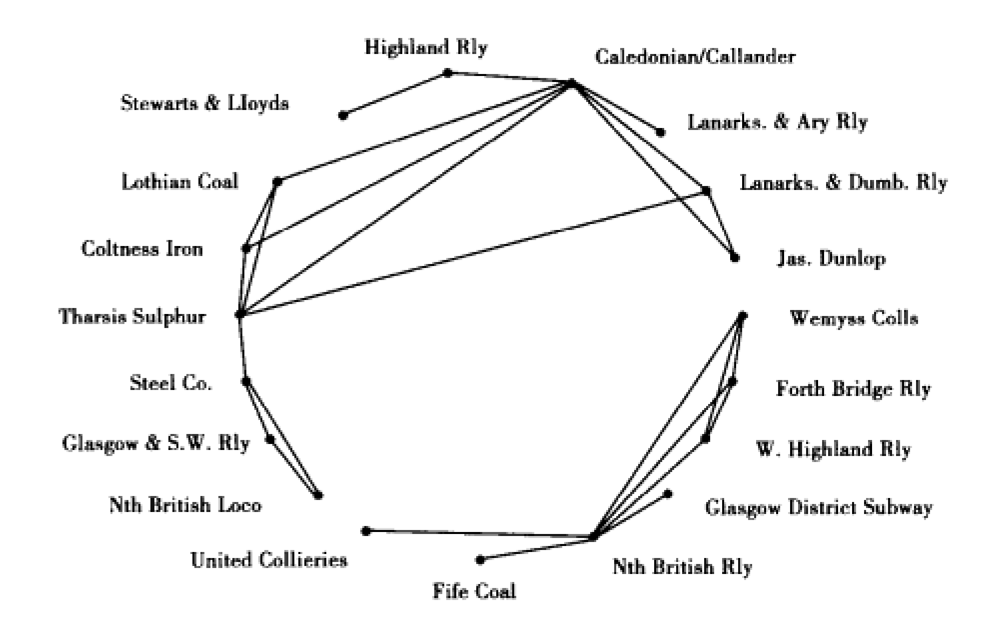


图 8.1 多家苏格兰公司:一个圆形图

排图示,就可以确定同一网络中数量不同的子群。这意味着在选择空间框架和确定其中点所处位置的标准的时候,都必须非常小心地对待。

麦格拉斯等学者的结论是,如果研究者希望推断出关于一个网络的实际的社会计量学特征的话,那么各个点之间的空间距离应该尽可能地与这些点之间的图论距离相吻合。这一点强化了社会网络分析者的一种长期期望,即走出隐喻的、说明性的图示,进而产生出更有说服力的社会结构图,从而像地理图一样可以在保持图的数学特征的同时,允许一些新的特征得以发现。这些图还有一个优点,即对于那些阅读研究报告的人来说,它们使得数据易于理解,并且更富有意义。

"多维量表(multidimensional scaling,缩写为 MDS)"这种数学研究体现了常规社群图及其各种推广形式(如圆形图)的全部优点,但却导致与网络所嵌入的空间"图"非常接近的结果。这是非常重要的一个进展,这种进展回应了场论的一些核心洞见。例如,正如不列颠诸岛的二维图可以允许其使用者做出关于该国的一些新发现一样——只要他们熟悉阅读图的一些原则,通过多维量表生成的社群图也能够带来关于所研究的网络的新知识。

距离、空间和量纲

多维量表(MDS)背后的基本思想是利用空间和距离概念来画关系数据图[2]。对于任何空间和距离模型来说,如果在其各个特征之间存在已知的、确定性的关系,那么可称该模型为一种"量纲(metric)"。就所研究的关系数据图来说,一种量纲框架具有一些有趣的特征。如果由多个点和线构成的构型(configuration)可以制作成为一个量纲图,那么就可能以不同于图论的方法来测量"距离"和"方向"。在图论中,两点之间的距离要根据连接两点的"途径"中的线数来测量①。因此,距离指的是"途径距离"。量纲式的距离概念与我们日常理解的物理距离概念很类似。例如,在一个"欧氏"量纲中,从A到B的距离完全等同于从B到A的距离,并且根据连接二者之间的最直接的路径来测量[3]。它是一种沿着路径的"直线距离",并且可以穿越"开放地带",而无需——确实也用不着——遵循一种图论意义上的途径。例如,在图8.2中,A和D之间的途径距离是3条线,而二者之间的欧氏距离为2厘米。

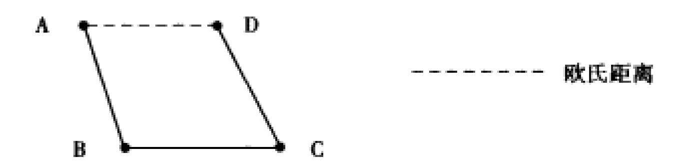


图8.2 欧氏距离

最简单地讲,多维量表试图把诸如途径距离这样的图论测度转换为与物理距离相似的量纲式的测度[4]。尽管我们对"欧几里德"这个术语可能不熟悉,并且该术语也确实令人畏惧,但是它却描述了我们最熟悉的、日常的距离和空间概念。因此,在社会网络分析中,该模型是一种特别常见的模型。关于各种社会关系的欧几里德图可以通过与地图册、地图集和其他我们熟悉的日常生活的空间模型进行对比来理解。

多维量表可以采用图论意义上的各个点之间的"接近度"测度,并且可

① 实际上,在图论意义上的两点之间的距离具体指的是连接这两点的"最短途径"中的线数——译注。

以用量纲式的术语来表示这些接近关系和距离关系。比较正式地讲，这涉及利用“相近性数据(proximity data)”来建构各个点的量纲式构型。在这种分析中，第一步是根据各个图论测度制作一个案例-案例相近性矩阵(case-by-case proximity matrix)。在这个矩阵中，每个单元的值显示出两点之间在多大程度上“相似”或“相异”。针对关系数据的相近度测度有几个，诸如个体之间交流的频率或数量、企业之间的持股关系规模、组织之间的共享成员数等。这些值中的许多量纲式特征将模糊不清，而且也很难知晓一个特定的测度是否同欧几里德式量纲中的一些假设相一致。恰恰由于这个原因，它们经常被转变为相关系数，因为我们知道相关系数与欧几里德量纲是一致的。在一个图中，有着相同关系模式的两点也将完全相关，所以这一对点的相近测度值为1。这种测度被称为相似性测度，因为其值越高，表示这两个点越“相近”。在这种情况下，即可认为相邻矩阵包含了“相似性”数据。另一类重要的相似性测度是“不相似性(dissimilarity)”，其值越低，表示两点越“相近”。对于研究者来说，他们必须清醒地认识到，自己是运用“相似性”还是“不相似性”来进行特定的多维量表程序的操作，这一点是至关重要的，因为二者可能产生不同的结果[5]。不论使用哪类数据，其目的都是为了形成一种量纲式构型，在这种构型中，量纲式距离模式对应着相似性模式。

如果考虑把一个城镇网络制图成为一页二维地图，就可以很容易地展示多维量表的运作方式了。镇与镇之间的里程矩阵包含了相近性的里数测度，这种测度可以轻易地在地图册上转换为点与点之间的厘米级距离。这种“量度(scaling)”将给出一个二维构型，其中的距离是根据东西和南北维度来安排的。然而，这种地图可能与所测量的国家的实际城镇之间的距离不完全吻合。城镇之间的道路很少沿着最短和最直接途径来建设，因此，道路里程也不会是真正的欧氏距离。同样，地图也不会考虑到第三维——高度，实际的路途可能翻山越谷，而非穿过纯粹的平原。不管怎样，地图仍然能够达到对实际情况的一个比较合理而有用的近似，并且其与任何较佳的解之间的“拟合之差(lack of fit)”是可以评估的。

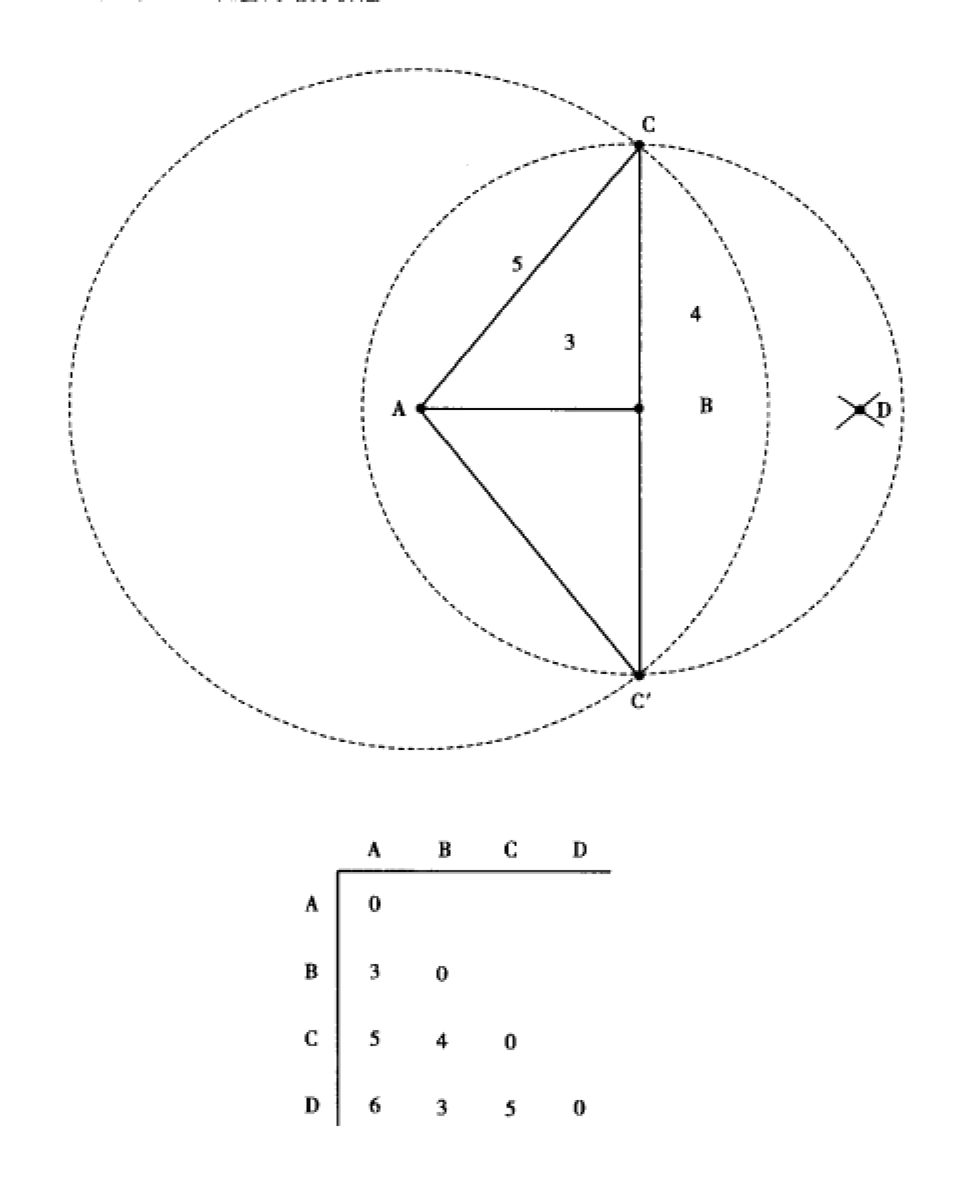
可见，一个简单地理图的构建可以很好地洞察到关系数据的多维量表输出结果。如果考察在构造图的时候使用的一些几何原则的话，就可以理解多维量表对关系数据的处理方法。我们可以通过一个非常简单的例子，即画一个三个城镇之间的距离关系图来观察这些几何原则，这仅仅需要根据它们之间的距离来构造这三个城镇的恰当的空间安排和位置。实际上，这项任务对应于一个传统学院派几何学中的一个经典问题，即在仅知道三边长度的情况下画一个三角形。

这个几何问题的解是把三角形的各个角作为圆点，各个点之间的距离作为半径来画圆。例如，考虑一个三角形，其三条边为 AB（长 3 厘米），BC（长 4 厘米）和 AC（长 5 厘米）。在构造这个三角形的时候，第一步是画出任意一边，如 AB 这条边。这条线可以画在纸上的任何位置。而我们知道 AC 长为 5 厘米，就可以推出 C 点肯定是位于以 A 为圆点、以 5 厘米为半径的圆周上。同理可知，C 点也一定位于以 B 为圆点、以 4 厘米为半径的圆周上。因此，构造三角形的第二步就是画这两个圆，确定交点。C 就位于交叉点上，因为只有该点才符合对这三条线的距离要求。实际上，如图 8.3 所示，交点有两个，因此，C 点的位置也可能有两个。此时，只选择其中一点作为 C 点就足够了。至于选哪个点并不重要，因为三角形 ABC′无非是三角形 ABC 的镜像。这就是对三个镇进行画图问题的解。如果 A、B、C 代表城镇，已知 AB、BC 和 AC 的长度代表镇与镇之间的标示里数，那么三角形 ABC 就是一个关于这些城镇位置的简单二维图。

但是，假如我们把 C′点而不是把 C 点作为第三个城镇的位置会怎样呢？ABC′可以同样被认可是一幅图吗？三角形 ABC′只是原来图形的一个倒置而已，所以无需在选择 ABC 和 ABC′中的哪一个作为城镇图方面犹豫。当一个构型完全是另一个构型的“映射”时，选择哪一个都无关紧要。哪个构型最有用，这只取决于地图使用者的方便，一个实际图与其映象图包含的信息完全相同。

同样的几何程序也可用于对四个或更多点的分析上。如果已知点 D 与 B 相距 3 厘米，与 A 相距 6 厘米，与 C 相距 5 厘米，那么它的位置就可通过分别以 A、B、C 为圆点所画的圆周交叉点来确定。一旦点 C 的最初位置得到确定，那么 D 点就唯一由这些圆的交叉点决定（参见图 8.3）。如此看来，作为一个一般性原则，只要最初三点的位置已经固定，就存在一个独特的二维构型图[6]。

可以把这种学院几何（school geometry）看成是为一个量纲空间中一系列点的定位提供了一个二维的解。在图 8.3 所展示的距离矩阵中可以构造一幅二维结构图。这两个维度分别是纸张上常规的横轴（左右）和纵轴（上下）。在为图集制作一张地图的时候，一般把各个点的构型向一定的位置移动，使得最北部的点指向页面的上方，最西部的点指向页面的左方。这样，横轴和纵轴就代表了已知的东/西维度和南/北维度。在多维量表中，这种构型的移动称为“旋转（rotation）”。然而，只有在常规性绘图中，旋转才主要涉及到把构型按照与已知维度相一致的方式进行排列。就多维量表分析的典型结果而言，旋转的主要目的在于发现一些有意义的维度。笔者将在本章后面回过头来再讨论旋转问题。



	A	B	C	D
A	0			
B	3	0		
C	5	4	0	
D	6	3	5	0

图 8.3　量纲模型的建构

上文描述了在一个简单的几何案例中所采用的方法，实际上，多维量表遵循的方法与之类似。虽然用来分析多维量表的电脑程序一般不运用这种不充分的办法来建构圈，但结果是一样的。多维量表的最初形式来源于托格森（Torgerson，1952）关于心理量表的开创性著述。如笔者所言，量纲性多维量表把相近性数据作为欧几里德距离来使用。如道路里程一例所示，如果使用不相似性测度，那么与我所描述的几何原则类似的一些原则就可用来制作一张量纲数据图。此时把原始数据看成是距离测度，并按照比例最终绘制成图。多维量表的算法运用一些几何原理来确保已知的

相似性数据与图中点的最终构型之间的"一致"。量纲式多维量表的多种研究在方法的细节上存在重大区别，即在把该程序推广到三维或者多维的过程中，所用的方法是不同的。

主成分和因子

有一种与这种量纲式的多维量表很相近的研究方法，这就是主成分分析法（principal components analysis，简写为 PCA）。尽管 PCA 与经典的因子分析法之间存在重要的区别，但在有些文本中，主成分分析法也被称为"因子分析法"。但是，就目前的目的来说，这个区别不是特别重要，并且在 SPSS 中，主成分分析作为因子分析程序的一个部分是常见的，这也使得其成为得到广泛应用的一种数据分析方法（Daultrey，1976；Goddard and Kirby，1976；Kim and Mueller，1978；Kline，1994）[7]。这种技术来源于早期的态度及智力量表法，这些早期研究者试图在大量特定的成就或态度测度中找到根本的共同因子。比如"常人智力（general intelligence）"就被看作是许多关于逻辑推理能力的专项测试表现根本的一个"因子"。推而广之，在任何已知的数据库背后都有可能存在两个或更多不同的因子。主成分分析是作为一种分析案例-变量（a case-by-variable）矩阵的方法发展起来的，其目的是为了发现多个变量中共享的一个或多个因子或成分。它试图利用原始数据发现一系列坐标或轴（因子或维度），从而可以用来绘制出数据散布图。当利用一个案例-隶属项（a case-by-affiliation）矩阵中的关系数据的时候，所得到的散布图就具有如下形式，即从一个点到另一个点的空间距离和环绕方向传达了有关它们之间的相对位置的一些实际信息。

如果将主成分分析法运用于有关属性数据的案例-变量矩阵当中，我们就可以很轻易地理解这种方法。一个简单的 PCA 算法将首先把案例-变量矩阵转换为变量-变量（variable-by-variable）矩阵，从而表明了变量（即初始阵的各列）之间的相关关系。因此，新的矩阵就展示了各个变量之间的相关性程度。下一步就是搜索这个矩阵，找到那些高度相关的变量，并且用一个构造出来的人工变量代替，以此来测量各个变量之间的相关关系。如此看来，一组相互关联或者相关程度高于一个特定值的变量将用一个构造出来的变量来替代。这个构造出来的变量就称为第一主成分。再下一步是寻找另外一系列高度相关，却与第一系列变量不相关的变量。用来代替这些变量的、构造出来的新变量就是第二主成分。主成分分析法将持续

此类过程,其目的是确定一系列互不相关的主成分,它们加在一起,共同解释了在数据中发现的全部变差。这种全面的分析将一直持续到确定出所有可能的成分。通过这种过程,初始的变量-变量相关系数矩阵实际上就转换成为一个变量-成分相关矩阵。

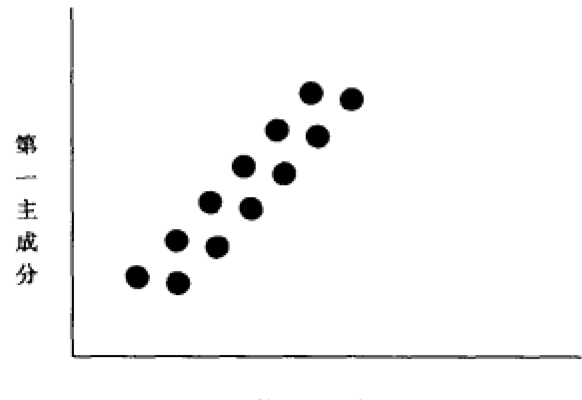
第一主成分代表了相关度最高的一组变量。根据定义,第二主成分与第一主成分无关,即与之是"正交"的。这种相互独立意味着,这两个主成分互成直角,可作为一个二维散点图的两个轴。这一普遍的原理同样也适用于数目更多的成分,其中每个维度与其他维度彼此垂直或者不相关。当然,画出或者图式出三维散点图比较难,实际上,维度超过 3 的散点图是画不出来的。不过,不管区分出的主成分有多少个,其研究的逻辑是相同的。在主成分分析中,一般需要搜索出能够解释数据中较高比例方差的最少数目的主成分。实际上,在任何一个停止的点都不能完全解释方差,这样的点也是随机的。对于研究者来说,如果新的主成分不能给已经解释的方差带来更多的贡献的话,研究者通常即可就此止步,不必继续分解新的成分。

因此,PCA 开始于一个变量-变量相关系数矩阵,它是根据初始的案例-变量矩阵构造出来的。根据变量-变量矩阵可以构造出一个变量-成分矩阵,该矩阵的每个格值表达的是计算出来的每个变量在每个成分上的"负载值(loadings)"。各个主成分被当成一个散点图的各个轴,各个负载值用来确定每个变量在各个轴上的位置。

旋转的可能性带来了一个复杂的问题,这一点已经在前文的简单几何例子中提到过。对一种构型进行旋转的目的是为了给它的结构提供一个更清晰的图像。例如,如果各个点在空间中似乎沿着一个特定方向散布的话,那么转动这个构型,直到各个点沿着第一主成分最大限度地展开,这将是合理的。例如,在图 8.4 中,图(2)表达了一种构型的旋转,这要比未旋转的图(1)更与轴拟合。更一般地说,旋转程序的目的是为了形成一个构型的定位,以便更好地与主轴吻合。旋转程序输出的结果是一个新的变量-成分矩阵,它包含了每个变量的一组修改过的负载值。

我已经介绍了 PCA 的一般步骤,它的结果是变量在由一组维度界定的一个空间上给出的散点。按照类似的方法可采取另外一个程序,该程序会形成案例的散点图。在这个程序中,第一步是把初始的个案-变量矩阵进行转置,并计算出一个个案-个案的相关系数矩阵。随后的程序则完全相同,尽管分析单位变为个案而不再是变量。最终结果是一个个案-成分负载值矩阵,从中可以绘制出个案分布图来。在某些文献中,这种个案-个案分析称为"Q-模"主成分分析法,以便与常规性的、变量-变量的"R-模"分析区分开来[8]。

（1）未旋转的解



（2）旋转后的解

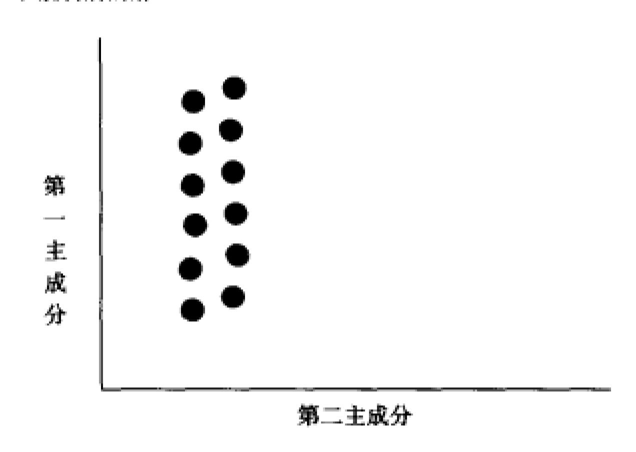


图 8.4 旋转

在处理属性数据的时候，我们需要区分 Q-模和 R-模。我们必须记住，通常情况下，R-模的 PCA 是针对初始阵的各列进行操作的。因此，对一个个案-隶属项发生阵（case-by-affiliation incidence matrix）的社会计量考察会导致对各个隶属项的分析。如果研究者希望考察各个个案之间的关系结构，就有必要对这个矩阵实行转置，从而把个案变为矩阵的各个列，如在一种 Q-模的 PCA 分析中一样。例如，可以对一个个人-组织发生阵直接进行分析，从而形成一个组织-成分矩阵，或者把这个矩阵进行转置分析，从而形成一个个体-成分矩阵。

在对这个“个人-组织发生阵”的直接分析中，PCA 算法可以考察各个组织，以便找到那些在其成员属性上最相似的组织。如在 CONCOR 中一样，此时也用相关系数作为相似性测度。所发现的各个具有相似性的组织

集合将用一些主成分来代替，这些组织可以在由各个轴定义的空间中像各个点一样画出来，而轴与成分是对应的。在这个散点图中，两点之间的欧氏距离可用来作为测量组织之间相似性的测度。对初始发生阵实行转置就可以分析各个个案。如前文几章所讨论过的 CONCOR 和块模型程序中所看到的那样，从同样的输入数据中产生的列解（column solution）和行解（row solution）将相互具有“对偶性（dual）”。它们虽然不同，但却是对同一组数据中的互补性表征。

尽管对于无向关系的对称数据矩阵来说，行解和列解将完全相同，但是仍然可以用 PCA 分析邻接矩阵。然而，在一个有向邻接矩阵的情况下，两种解会有所不同，一个解可能对应于由“发送”关系生成的网络，另一个解则对应于由“接受”关系生成的网络[9]。

一些非量纲的方法

在利用量纲式多维量表和主成分分析法来探究关系数据的过程中存在很多局限。许多关系数据在形式上是二元的，仅仅表明关系的有无。这种数据就不能直接用来测量相似性。正如笔者所指出的那样，二元数据必须首先转换为具有量纲性质的某些测度，如相关系数等。但是这种数据转换过程可能导致研究者做出关于关系数据的一些未加证实的假设。即使初始数据为多值数据，一些量纲式的假设也可能是不恰当的。特别需要指出的是，使用定距或定比测度就可能是不恰当的。例如，对于拥有四个共同董事的两个公司来说，我们不能认为这两个公司之间的关系要比仅仅有两个共同董事的两个公司之间的关系密切“两倍”。尽管从现实角度讲，可以认为前一种关系比后一种关系更“密切”一些，但是很难确切地知道密切的“程度”多大。幸运的是，人们设计出一些功能强大的多维量表技术，它们不要求直接输入量纲式数据，并且其应用范围要比对应的量纲式方法更广泛[10]。

这些非量纲式多维量表技术也常常被称为“最小空间分析（smallest space analysis）”，它在标准的计算机程序中得到了广泛采用。非量纲式多维量表程序的操作基础是一个对称邻接矩阵，其中每个格值表示各个案例之间的相似性或者不相似性，表达方法既可以利用相关系数，也可利用实际多值数据。该程序不把这些值直接转换为欧氏距离，而是仅仅考虑到这些值的等级次序。也就是说，这些数据是按照等级层次来对待的。非量纲

式程序追求一种解，在这个解中，各个距离的等级次序与初始值的等级次序是相同的。

利用图8.5的数据就可以说明这个程序。第一步把初始阵中的各个单元值按照降序（从高到低的顺序）排列，从而构造出一个新的矩阵，在该矩阵中，各个经过排序分布后的等级值代替了初始单元值。例如，在图8.5中，初始阵中A与B的不相似程度最高，因此就在等级数据矩阵中用1来代替；A与C之间的不相似程度是6个值中的最小的，因此，在等级数据矩阵中用6来代替。这样看来，很有必要建构一个欧氏距离矩阵，这些距离像初始单元值一样有着同样的等级排序。

（1）初始矩阵(不相似性)

	A	B	C	D
A	0			
B	60	0		
C	10	40	0	
D	30	50	20	0

（2）等级数据矩阵

	A	B	C	D
A	—			
B	1	—		
C	6	3	—	
D	4	2	5	—

（3）负载：二维的解

	维度1	维度2
A	0.575	0.404
B	-0.993	0.114
C	0.195	0.177
D	0.222	-0.695

图8.5 非量纲式多维量表数据

可以用这些欧氏距离绘制一个与PCA生成的图类似的量纲式散点图。在这种情况下，距离的等级次序就和相似性的等级次序一样，但是此时却用不着做出关于相似性数据本身性质的一些假设。如果相似性测度确实具有一些量纲性质，该程序当然会形成一个最终矩阵，如果只考虑全部值以同样的数量递减的话，那么其中的每个值都和初始阵的那些值完全对应。图8.5中的例子不属于此情况，因为初始数据不具有量纲性。尽管一种简易几何方法的变体不能应用于初始数据，却可用于最终矩阵的每个格值上。最终矩阵表达了量纲距离与非量纲数据之间的“最佳拟合”，第（3）

组表示的是4个点在二维解中两个维度上的负载值。

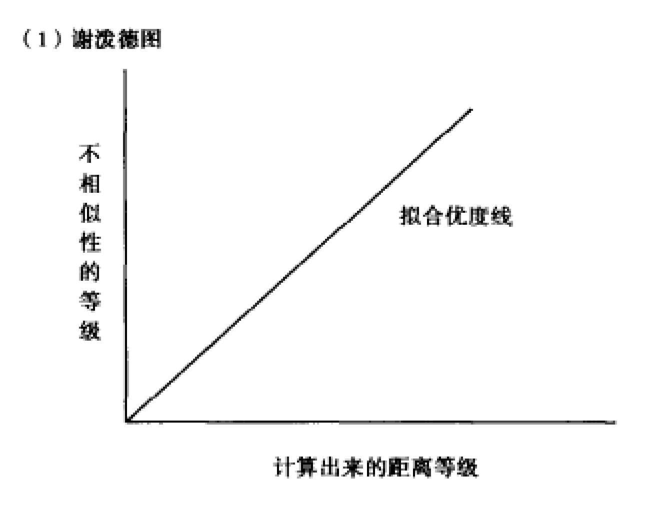
接下来的问题是,如何计算欧氏距离矩阵?通常的算法开始于计算一种猜测值,即猜测出这些距离会成为什么样的。这个猜测形成了初始的或者试探性的构型,该构型的各个距离的等级秩序要与相似性构型中的等级秩序进行比较。对初始估计值进行的后续"试错(trial-and-error)"性修正,会逐渐导向更好的试探性构型。如果距离等级排列和相似性等级排列之间有较好的吻合的话,那么就认为这个构型是对此前构型的一种推进,从而接受它。最后将会发现一种与初始数据达到最佳拟合的试探性构型。

因此,为了开始这个过程,必须建构一个恰当的试探性构型。初始构型本身可随机生成,或者,如果已知网络的某种结构,研究者也可以提供距离的估计值。方法的选择并不重要,因为初始构型只是分析的起点,它精确与否不影响分析的其他方面。利用随机选择的起点的唯一缺点是,分析过程可能会稍长一些,因为在发现最终的构型之前需要经过很多步骤。实际上,得到广泛使用的 MINISSA 算法通常会通过主成分分析来建构一个初始构型[11]。

怎样才能知道产生的最终构型是令人满意的?在初始构型中各个距离的等级排序必须与最初的相似性数据中的等级排序进行比较,以便发现每一对点在等级上出现什么差异。这种比较展示了各个点必须在哪个方向上以及在多大程度上进行相对的移动,以便减少这种差异。例如,如果在等级上确实存在显著的差异,那么各个点必须要比试探性构型更向相互接近的方向移动。对差异性的这种比较要在每种后续构型中进行。每一步的结果都可以画成如图8.6(1)所示的图。在这个图——其取名来自该图的提出者谢泼德(Shepard)——中,在试探性构型中各个距离的等级排序要根据最初数据中的各个不相似性值的等级排序来画出。如果各个点广泛地分散在图中,则拟合较差;如果它们聚集在45度线左右的话,拟合度就较好。只有当各个点都在45度线上的时候,才达到最佳拟合态[12]。持续地加以调整以降低不一致性,就会逐渐产生一种构型,在该构型中不能做任何不削弱这种拟合度的改动。当达到这一点时(这也许要经过相当多的步骤),就最终达到与最初数据的最佳可能的拟合。

行文至此,维度数似乎是不成问题的。确实,在图8.5中,这个简单示例假定的是一种二维度的解。在实践中,研究者必须决定用来绘制一份数据图时所需要的维度数。

对每个维度的解来说,都可以画出谢泼德图(Shepard diagram),并且可以发现一种"最佳拟合"的构型。例如,在二个维度上将有一个最佳拟合,在三维上也将有一个最佳拟合。但谢泼德图并没有指出哪种解更好。



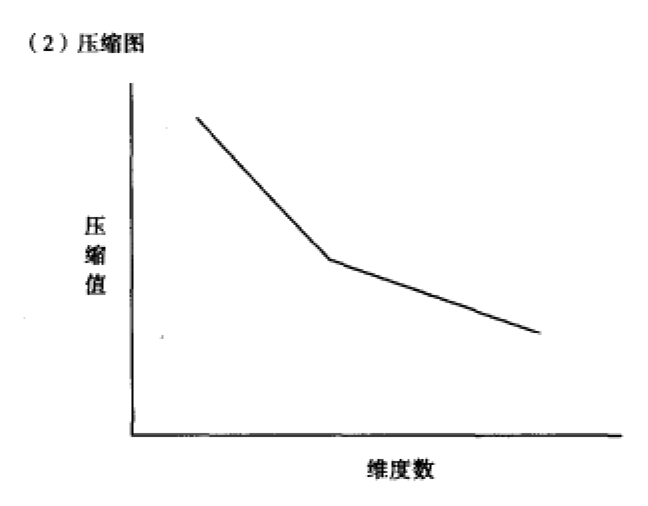


图 8.6 在非量纲式多维量表中的拟合优度

非量纲式多维量表法本身并不决定维度的合适数量是多少。研究者必须针对不同数量的维度进行大量的分析,试着发现在各种维度解中,哪个给出了与初始数据的最佳拟合。这一点可以通过计算一个名为"压缩值(stress)"的统计量来获得。该统计量测量了谢泼德图中各个点围绕拟合优度线的平均分布情况。通过绘制每个维度解的压缩值,就可以画出一个类似于图 8.6 的图。由该图可见,压缩值最初会随着维度数量的增加而减少,但是最终会达到一个"拐"点,即随着维度数的进一步增加,压缩值的降低就不那么明显了。一旦达到这种平稳状态,就达到了最可能的维度解。维度数再增加,可能降低压缩值,但没有达到可感知的程度。

除了要比较各个压缩值之外,也有必要考虑到在优选解中压缩值的绝

对值。克鲁斯卡尔(Kruskal,1964a)认为,如果任何解都不能将压缩值降低10%,那么结果就不能被认为是对初始数据的充分拟合。克鲁斯拉尔建议,下降5%或更低表明拟合程度好,而10%仍然是比较“适当”的,但是接近20%的压缩就是“糟糕的”。在图8.5中,数据的二维解的压缩值就是0。

在量纲空间中,“维度”观念并不是社会网络分析中所提到的唯一维度性(dimensionality)概念。阿特金(Atkin,1974)已经在Q-分析法基础上提出一个维度性的概念。在Q-分析法中,一个点的维度性比发生阵中的行或者列的解小1。根据阿特金的观点,这个值给出的就是能够充分代表一个点所需要的维度数。因此,例如,一身兼四个公司董事的人必须用三个维度来展示。然而,从这个观点看,网络中的每个案例都有自己的维度性,这些可能不同于整个网络的维度性。

尽管这种研究的优点在于它根源于与图论很接近的一些观念,但是它相对于我们熟悉的一些概念来说,其价值不明显。因此,弗里曼(Freeman,1983)拒绝了针对图的维度性的Q-分析观,而提议把图论和通过多维量表发现的几何维度结合在一起。他指出,一个图的维度性就是把图高度拟合地嵌入一个空间时所必需的最小维度数。然而,他提出的拟合优度标准要比克鲁斯拉尔的标准更强,并且包括了压缩值为0的情况。

在一张平面纸上画一个二维散点图,这是我们熟悉的,也易于理解,但是,在完全绝对水平上给出的最佳拟合的维度数经常大于2。在这些情况下,就不可能在一张平面纸上画出最终的构型图来。可以用透视法或者制图技术画出第三维,但是这种画法几乎没有什么说明性(参看Levine,1972,下文有讨论)。对铅笔和图纸法的重大改进便是利用电脑绘图法绘制出一个三维构型图,但是如果运用三个以上的维度的话,这种程序在运行时也会遇到类似的问题。对于维度数目多的情况来说,最常见的解便是在纸上依次展示出穿过构型的各个二维“片”。例如,在一个三维解中,可以在纸张上把这个构型表征为三个独立的在二维情形下看到的整个构型:维度1和维度2情形,维度2和维度3情形,维度1和维度3情形。

运用来自多维量表程序的输出项,各个点的构型可以绘制在一个空间内,这个空间是由所发现的维度的数量界定的。随后就可以开始解释的过程。有两个问题需要解释:维度的意义和各个点在空间排列上的重要性。例如,在地图册中的一张地图上,一般无争议地把维度看成是指常规地理空间上的东/西、南/北两个维度。然而,对于社会网络模型来说,分析者的最初任务就是达到对各个维度作出的社会学解释。比如,可以确定一个维度反映了个体的经济来源,而另外一个维度反映的则是他们的政治归属。

对构型进行旋转常常有助于解释各个维度。对各个点自身的空间分布赋予某种意义也是有必要的。一个常规的程序就是利用对原始数据进行聚类分析时得到的输出信息，把各个点进行分组。各个聚类内的点由轮廓线围绕着，利用等级聚类法就可以构建出各个点的轮廓图。例如，可以利用伯特提出的一类对欧氏距离进行的聚类分析法来绘制出一个欧氏多维量表解的各个聚类分布图。另一方面，对图的各个成分和核心进行的以多元度和度数为基础的分析结果可以绘制在一个多维量表的解上。有鉴于第六章讨论的嵌套成分的"轮廓图"可以随机地画在纸上的各个位置上，那些与一个多维量表的解相关的轮廓图则可以代表一种更"自然"的数据排列。图 8.7 就给出了这种程序的一个简单例子。

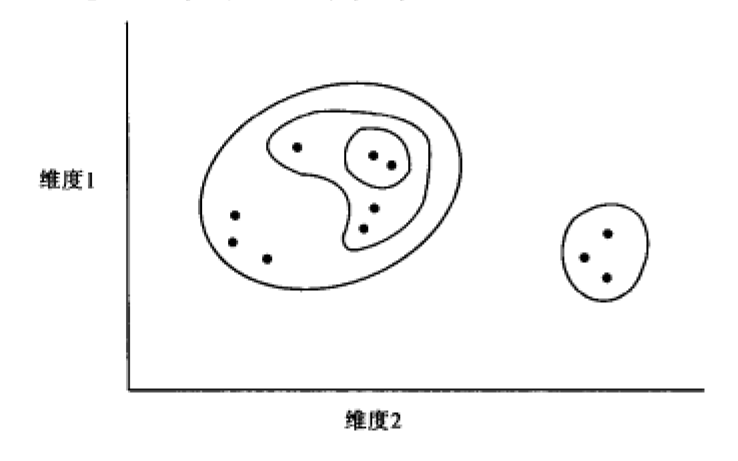


图 8.7　多维量表构型图

在画出如图 8.7 所示的一个图之后，研究者就可以搜索出聚类中各个点的某种共同特征，因为这些点是该空间中的一些独特邻点。这样，一个完整的解释就涉及要确定各个聚类，然后利用一些维度对各个点的相互位置赋予一定的意义。应该注意的是，这种解释的过程对研究者来说是一个创造性的、富于想象力的行动，它不是仅仅通过电脑就能生成的。研究者仍然要控制这个过程，并且要对赋予结论的各种解释负责。确实，不能保证各个维度能够提供什么实质性的解释；它们可能只是抽样或分析方法的人造物。例如在心理学中，关于智商（IQ）的测试是真能反应普遍智力中的一个真实因子，还是无非再现了测试过程的特征而已，这个问题就有很多的争议。

在网络可视化方面的一些进展

多维量表以及与之相关的一些技术在把一些图论概念和空间观念整合在一起方面做了大量的工作。尽管如此，这些分析的结果——特别对于巨型社会网络来说——常常仅仅表现为点和线的稠密集合，而该集合不能轻易地进行可视化考察。因此，许多社会网络分析学者已经探讨把多维量表与一些有效的结构建模技术结合在一起的可能性，这些技术有助于以更直观的方式来探讨网络结构并使之可视化。

进行此项工作的最早尝试者之一是克劳夫达尔（Klovdahl，1981，1986，1989），他的 VIEW-NET 程序运用了分子建模方法，致力于利用简单的三维球-棒关系来表达点和线。然而，对于大网络来说，该程序的输出结果仍然不易于解释。

为了超越这种局限，克雷姆珀尔（Krempel，1994）提出一些把复杂结构进行简化的方法，进而突显它们的本质特征。他的方法要利用一种简单的——常常也是任选的——几何形作为框架，以此来对一个网络的各个特征进行组织。当然，图 8.1 所展现的圆形图只是这个程序的一个简单例子罢了，克雷姆珀尔也把圆作为他自己研究的基础。他又进一步推广这种研究，设计了一种算法，该算法利用一些图论测度来生成关系数据与圆形之间的最佳拟合——恰恰像一条回归线提供了与点的散点图的最佳线性拟合一样。因此，距离测度或中心度测度可用作决定环绕圆的各个点的位置的审美标准。只要数据或多或少是由不同的子群体构成的，就可以把这些子群体分析成为处于一个大圆中的各个独立圆。我们头脑中已知的圆形为我们提供了一个解释实际网络的各个总体特征的熟悉结构[13]。克雷姆珀尔认为这个程序能够产生某些特殊关系构型的“潜在结构”。

这个程序也为处理大规模数据集合提供了极大的可能。我们有可能把分析得到的低层次的圆压缩成能够代表整个子图的一些宏观点（macro-points）。然后可以生成一个有关宏观点（macro-points）的社群图，从而提供一个简化的、清晰的模型。研究者可以选择该社群图的任意一个特定部分作深入分析，也可以解压一个宏观点，以便分析其内部结构。这样看来，一个克雷姆珀尔图（Krempel diagram）就是由各种嵌套的圆构成的等级体系，其展示的细节数量依赖于哪些圆得到了压缩，哪些圆又得到了解压。

弗里曼（Freeman，1966a）试图给出网络可视化的一些普遍原则，他设

计了一些物理学方法，并将它们用于社会网络分析。他指出，点和线至少应被进行着色编码或者被区分开来，以便显示出它们的关系性质，其设计也应该围绕着数据中的一些最重要的结构特性加以组织，并且一个网络中的依赖于时间方面的要素也应该通过一些动画程序展示出来。

弗里曼特别感兴趣的是把化学中的分子建模技术应用于分析社会网络数据。在这些技术中，点和线用三维的球和棒来表达，并且现有的大量计算机程序可执行这种操作，而不用物理模型来展示。他大量用了 MOVIEMOL①，这是一种动画程序，可用来展示（化学或者社会）结构伴随时间而发生的动态演变。然而，这种程序限制了点的定位，因为它假定化学法则应该决定各个点的定位。尽管弗里曼（Freeman，1996b）发现该程序在分析小网络的变化方面是有用的，但是到目前为止，很难把它修正为分析社会网络数据的程序。

一个更有灵活性的建模程序是 MAGE[14]。该程序在安排点的位置方面有更大的灵活性（例如该程序可利用多维量表的输出文件），并且可根据线的强度或多元度来画线。通过把这种程序与基于网络的虚拟现实建模语言（VRML）②可视程序（viewers）（它可以很容易地与诸如 Internet Explorer 和 Netscape 这样的网络浏览器整合或叠加在一起）链接在一起，就可能围绕网络模型的多个维度对它进行旋转，并通过扩大和缩小方法来更细致地探讨该网络模型。尽管 MAGE 是以一系列静态而非连续流的方式来运行的，但是它也可用来展示随着时间的推移而出现的一些变化。未来的发展可能涉及利用一些形态学程序（morphing procedures）来得到一些真正动态性的效果。

精英，社区和影响力

在社会学中，多维量表的最早应用之一是由爱德华·劳曼（Edward Laumann，1966）于 1960 年完成的社区研究。该研究通过随后的一系列有关社区权力和精英结构的相关研究而得到了发展和推广。这项研究成果发表在劳曼在哈佛大学的博士学位论文中，该论文是在霍曼斯指导下完成

① Moviemol 是一种简单易用的分子、原子、离子结构的可视化动画程序，这个软件生成的动画文件为 MPG 格式，但文件较小，主要用于模拟微观世界的分子运动——译注。

② VRML 为 Virtual Reality Modeling Language 的缩写——译注。

的，也显示出受到了帕森斯和哈里森·怀特的影响。这项研究把帕森斯式系统理论的一般理论框架和哈佛学者的高级社会计量学研究结合在一起，并将这些方法和理论引入到社区结构研究中。然而，在进行数据分析时特别坚持的导向来源于路易斯·格特曼（Louis Guttman）的研究工作，而劳曼在密歇根曾与之谋面。正是由于受到格特曼的影响，劳曼才决定利用非量纲式的多维量表技术分析他的关系数据。

劳曼的研究开始于博格达斯（Bogardus）的“社会距离（social distance）”观念，该观念是在1920年代以来的各种文章之中（参见Bogardus，1959的总结）出现的。劳曼对社会距离思想的解释如下，即社会距离代表了在各种职业位置的占据者之间发现的不同关系模式。这是一种“客观的”测度，能够测量出社区生活中各种社会位置的占据者之间相互联系的程度有多大或者多小。劳曼把这种“客观的”社会距离概念与能动者在相互接受积极或消极态度之时可能体验到并表达出来的社会距离“主观”感受进行对比。如此看来，劳曼的研究工作严格地遵循如下思想传统，即将对社会计量学的关注从心理学（或者“主观”的层次）领域向关系性关联的社会学层次转移。他的目的是达到对“客观的”社会距离进行操作化测量，然后运用多维量表方法把它转化为一个社会结构量纲图。

劳曼在马萨诸塞州波士顿市的两个城区抽取一个男性白人样本，目的是在样本中获得高度的职业多样性。他进行的是一种本书最后一章讨论的位置研究，因为他的分析单位是各个职业位置而不是个体本身。每种职业位置的占据者要回答的问题有：择友、亲属、邻里等，把这些回答综合在一起，就可以得到每个位置的一些汇总测度值。

最初的分析包括五种职业类别：高级专业和大型商业事务、半专业和中等事务、小职员的小型商业事务、技工、半熟练工和非技工。在建构包括频次资料的大量位置-属性发生阵的时候，需要用到这五个位置，然后按照传统的列联表方法分析这些数据矩阵。此统计分析表明，择友大大受到职业平等性的限制，而其他社会关系则比较容易包含不同职业位置的人。

然而，劳曼的工作真正创新之处在于，应用多维量表来发现在不同的关系模式中是否内在地存在等级结构（Laumann，1966：第6章）。对职业等级的传统研究的基础是“声望”等级，即用对某些特定职业所占据的积极评价来测量他们的地位。劳曼放弃了这种研究思路，因为该思路依赖于一种“主观”的评价，而他力图运用社会距离的实际关系模式来建构一种等级结构。他从邓肯（Duncan，1961）目录中选用了55个职业位置，建构成为一个55×55的发生矩阵，或者针对七种社会关系的每一种来构建“联合发生（joint occurrence）”矩阵。把各个独立的矩阵汇总成为一个表达不同关系

的发生阵，并把发生阵中的标准化频数值看成是各个相近性测度值。一个职业地位的成员与另一个职位的成员之间的互动频次越大，他们在社会空间中越“接近”[15]。

劳曼的分析结果表明，三维解与原始数据达到最佳拟合。在一个三维空间中画出55个职业位置，并且在邓肯声望指数上相互接近的数据周围画出各条轮廓线（Laumann，1966：图6.3给出了该构型的一个褶页图）。关于这种具有相当任意性的聚类的细节信息很少，这似乎恰恰内建于劳曼试图规避的“声望”假设之中。不管怎样，他对模型的解释仍然是有价值的。他把第一个同时也是最重要的维度看成是声望维度。这个维度上的分数与邓肯指数（Duncan index）之间的相关系数为0.824。确实，似乎不同关系模式确实遵循的是早期的声望研究所描述的模式。但是该模式不能根据简单的一维概念来解释。然而，其他二个维度不容易解释，并且劳曼也没有对第二个维度给出满意的解释。他暂时性地把第三个维度看作为创业者职位与科层制的有薪职位的对比。

后来，在劳曼于1966年进行的底特律（Detroit）研究中，他进一步推广了这种社区结构研究（Laumann，1973）。劳曼的研究目的是用由1 013个白人男性构成的样本，探索存在于各种社会位置之间的朋友关系。这项工作延续了早期研究中对位置的关注，但是它从职业位置扩展到其他社会位置。虽然这种研究方式与沃纳的开创性的位置研究有很多共通之处，但帕森斯（Parsons，1951）却是他所特别参考的理论点。

劳曼主要分析了有关友谊的族群-宗教和职业关系网数据。他最初分别分析了族群性和宗教，但却发现，如果把它们合成一个单一关系的话，就会得到一个更好的多维量表的解。22个族群和15个宗教群体被组合成27个族群-宗教群体，并根据他们成员之间的“择友”数据计算出各个不相似性测度。他给出的三维解（Laumann，1973：图3.3）显示出强大的第一维，该维度把新教徒（Protestants）、天主教徒（Catholics）和犹太教徒（Jews）分离开来。第二维似乎测量了经济地位，并且与家庭收入密切相关，而第三维测量了参与教堂活动的频次。因此，族群-宗教群体要根据宗教、收入和参与教堂活动这三个维度来构建。例如，把天主教徒分为高收入和低收入两组，并且各自独立地根据参与教堂活动的频数来区分。在社会空间中，可区分出来的职位聚类通常拥有一个族群基础。

在劳曼的研究中，职位分析包含16种职业组，研究发现，一个二维解达到最佳拟合。他对这些数据的讨论充分地肯定了前期研究的结果。他再次得出结论认为，第一维度是地位或声望（它与收入和教育水平相关），并且第二维度把创业性与官僚性区分开。

在与帕佩(Pappi)合作的时候,劳曼进一步把自己关于社区结构的分析推广到"精英"层次,这需要吸收亨特(Hunter,1953)和达尔(Dahl,1961)的工作以及他们对社区权力的研究(Laumann and Pappi,1973,1976)。他们研究了西德的一个名为竹力西(Julich)的小镇,并起了一个学名:奥纽斯塔德(Altneustadt)。在1950到1960年代之间,该镇是一个迅速扩张的城镇,并且在原有居民和迁入者之间出现了一些共有部门,这些部门分别拥有自己的政治关注点,即基督教民主同盟(Christian Democratic Union)和社会民主党(Social Democratic Party)。该研究在方法上仍然坚持位置导向。尽管对各个位置的个体占据者进行了抽样,但是重要的是各个位置之间的关系。他们关注的社会位置是在帕森斯式的A,G,I,L体制子系统中的最高权威位置。他访谈了46个这些位置的占有者,要求每个人指出其他45人中哪个人在小镇中最有影响力。在如下这一点上达到高度共识:"Herr K"作为最具影响力的人物赢得46票,因为在该位置上他也进行了自我推荐,显然,研究者必须认真地考虑矩阵的对角线了。劳曼和帕佩在三种社会关系中考察这46个人之间的互动:商业和专业关系、"社会的、表达性"的关系(例如根植于教育、宗教和居住地上的关系等)和社区的各种事务关系(政治团体和联盟)。

对每一种社会关系来说,都要画出社会计量学的选项图,但是,由于研究者仅仅对一种关系的有无(而不是方向)感兴趣,这些关系因而要转化为对称矩阵。在非量纲式的多维量表中需要用到捷径矩阵,而针对每一种关系都要解出各个二维解。对产生的结论性结构的解释是,在由各个点组成的构型中,可以把趋于其核心点的位置看成是在社区结构中具有"整合性的中心度(integrative centrality)"的指标。以"社区事务"网为例,该网络显示了社区的政治结构,该结构的中心区域由在权力结构中有影响人物构成的"内核"构成。图8.8就是劳曼和帕佩给出的竹力西小镇社区权力结构简化图。各个点被安排在中心度递减的各个区域之中,画出的各条政治部门线把那些对五个关键社区问题的每一个都持有相反意见的群体分开。把中心度和问题区域结合在一起,导致在社区权力结构中确认出了大量的不同部门。

有证据表明多维量表是一种有用的技术,可对社区权力的某些类型作进一步研究,进一步突现出地方精英的存在。给人印象极为深刻的是,该技术也被用于对国家层次精英的研究。劳曼和帕佩用声望方法来发现当地精英,而莱文(Levine,1972)却运用商业中的连锁董事关系来确认国家层次的经济精英。

在新一轮的哈佛研究者中,莱文是最初的网络学者之一,并且他开创

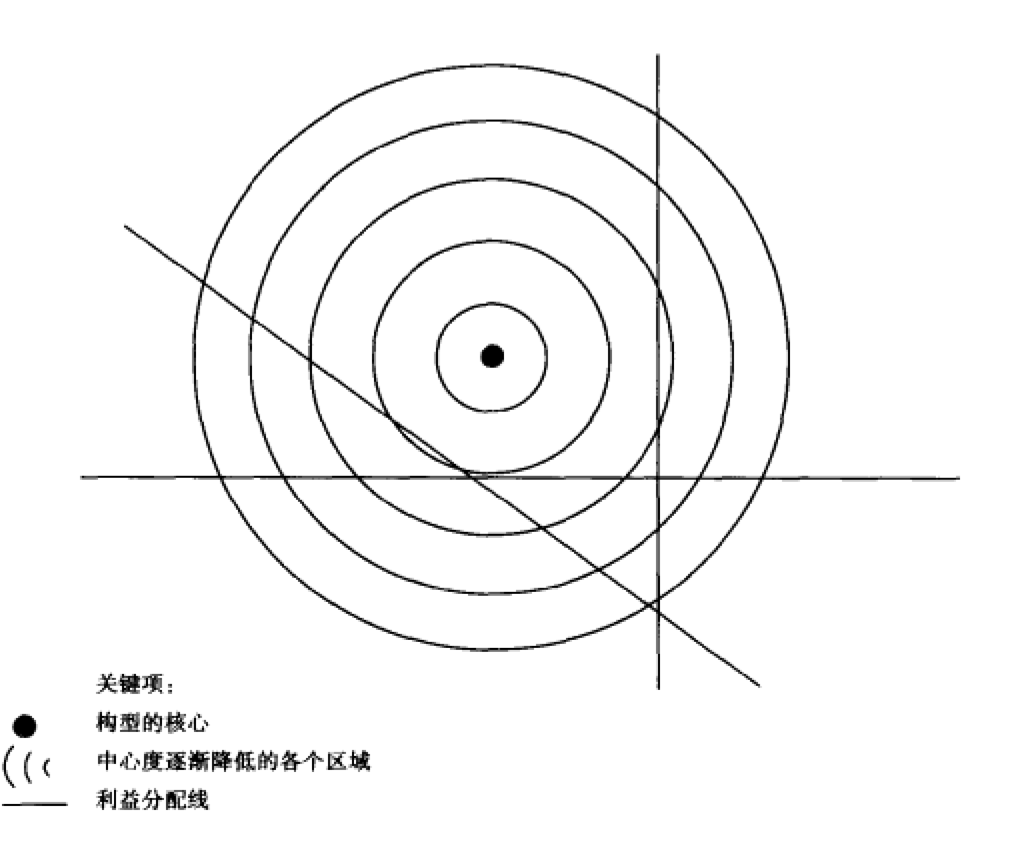


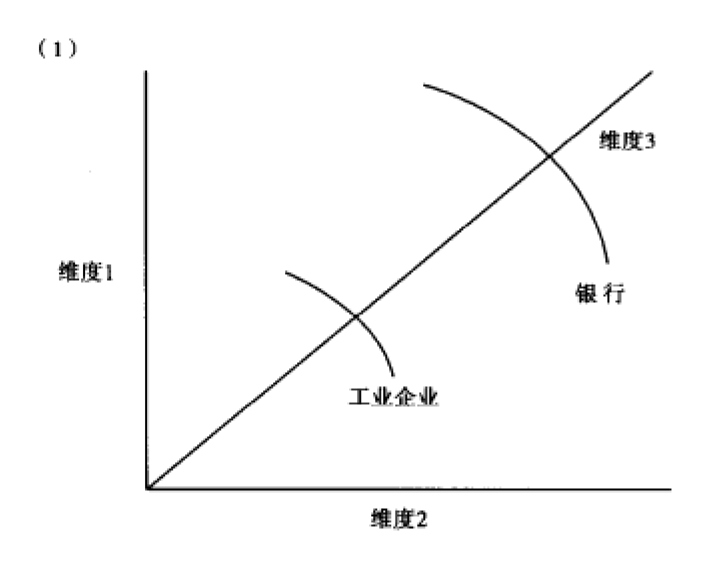
图 8.8 社区精英结构

性地运用多维量表（而不是图论）作为社会网络分析的技术。他运用 1966 年帕特曼（Patman）对美国银行运作的调查数据进行分析，探讨了 100 强工业公司以及它们与三个城市中的 14 家银行之间的关系。在这些企业中，共有 70 家与银行有连锁关系，所以，莱文构建了一个 70 × 14 发生阵，以便展示每一对企业共享的董事数目。共享的董事数被看成是公司之间的一种“偏好”或“相似性”测度。然后对发生阵进行分析，以便构建一个联合空间，在该空间中可对银行和公司进行定位。

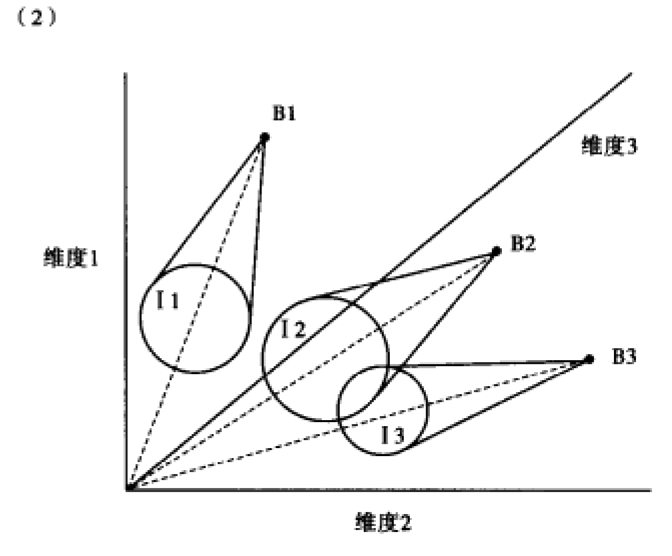
莱文的分析结果表明，三维解可以达到与初始数据的最佳拟合，莱文对其中的两个维度作出了如下社会学解释。首要的、最重要的维度显示出数据的一种区域性的结构，它把纽约的、匹兹堡的和芝加哥的企业分开。他对第二维度并没有给出解释，但是莱文认为，第三个维度把银行与企业分开。他的结论汇总在图 8.9 中。

莱文的观点是，可以认为结构形成了围绕联合空间之中心的一个球形域。这个球形域的结构是由类似于洋葱的分层排列或者同心圆序列组成的。沿着第三维可以看到，工业企业位于球形区域的不同内圈，而银行则

（1）



（2）



关键项：
B1，B2，B3：各家银行
I1，I2，I3：与每家银行有关的工业企业

图 8.9 多维量表和银行连锁

位于同一维度的外圈。这个模式反映在图 8.9(1)中。各个企业在图中各个壳形轮廓的确切位置由第一和第二维度来描述，这表明银行和工业企业都分别围绕各自的轮廓在地域上有所分化。

莱文认为，这种球形结构的中心位置将被一些孤立的企业，即那些不存在连锁的企业所占据。把这些企业从他的数据组中排除出去，在最终的构型中，中心就是空的[16]。与一些特定的银行（这些银行被马克思主义者描

述为金融利益集团)相联系的工业企业聚类可以被看成是一个球形区域中的各个部门或“楔子”,如图8.9(2)所示。如果从每家银行向球心画一条线,那么每条线或向量[17]就是该银行在圆形区域中的中心线。一个工业企业与银行之间的空间距离要沿着该向量进行测量,其值表明与银行之间的紧密程度,该向量与那条把工业公司与圆心连接在一起的线之间的夹角就测量了该工业企业在楔子中在多大程度上居于边缘地位。

莱文虽然把工业-银行连锁关系的整体构型描绘成一个球形,但他也把与每家银行联系的楔子描述成为它们的“影响域”。使用这些概念的理论基础在于,这两个术语与关于同一结构的不同观点相关联。如果从每家银行的立场出发来考虑,会发现有一些关系圆围绕着它们。如果从所形成的圆锥形楔子的顶点来看待其所连锁的企业(参见图8.9),银行会将这些企业看成是围绕它的圆形模式排列的。但是,这些银行的影响域会如此相互交织,以至于产生总的圆形的联合空间构型。从整个结构的角度看,银行领域似乎作为大的公司之间的关系域的圆锥楔子而存在。莱文把这种视角的差别等同于有关星系的地心说和日心说之间的差别。莱文并不关注特定银行的自我中心领域,而是转换视角,试图关注总体结构本身的社会中心特征。

莱文最后考虑的问题是如何在一个二维平面上最好地表达他的三维构型。图8.9运用透视画法技巧做到了这一点,但是这种方法的局限性已显现出来。为了提高他的展示力,莱文转向应用地理学家曾经采用的用一个平面来描述地球的三维结构的投影制图法。他拒绝运用任何由多维量表自身产生的二维观点,因为这些观点尤其会产生一些不希望看到的曲解。这种观点的基础是“平行投影法(parallel projection)”,它是一种来自于无限的观点。这个投影法很少被制图者使用,因为它会伴有失真的情况。图8.10展示了这一点,在该图中,一座山的山基和顶点在直线投影中是分离的两点。在平行投影中,顶点和基点在纵向的分裂造成在结果图中横向的分裂。

为了解决这个问题,莱文设计了一种来自中心点自身的“点投影”形式。在制图学中,这种形式被命名为日晷指针投影(gnomonic projection)。在这个投影中,所有处在同一条半径上的点被绘在页面的同一位置上。因此,如图8.10所示,一座山峰的顶点和基点是作为独立的点出现的。对莱文数据的充分表达将清晰地展示出一些特定银行和与之相关的工业企业之间的关系——它们应该在页面上作为一些明确的聚类而出现。莱文论证道,一个日晷指针投影将满足这个要求。每家银行出现在其聚类的绝对中心部分,而在圆内的工业企业更接近于与之相关联的、处在外圈上的银

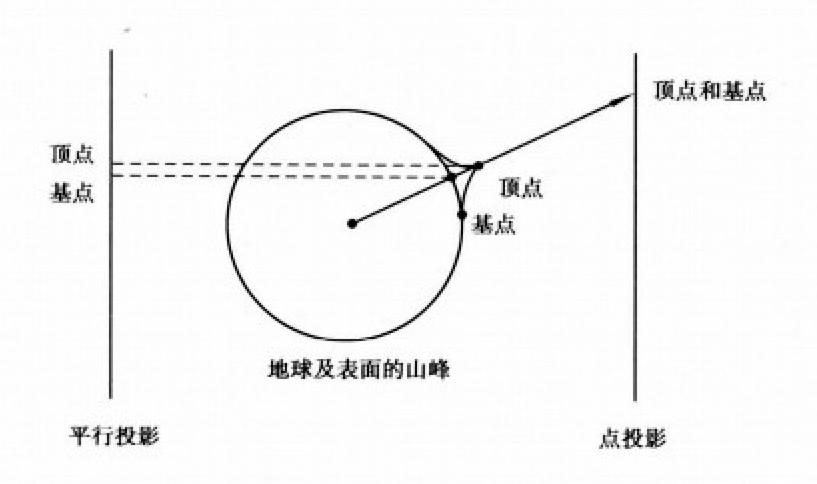


图 8.10 可选择的两种投影

行。利用这种方式,各个独立的"银行影响力范围"仍然保留在整个结构图中。该构型的日晷指针投影如莱文(Levine,1972:图 10)所示。

莱文的研究很重要,它标志着社会网络分析方法的前景和成就。它是由哈佛大学网络分析新小组的第一代学者做出的开创性研究之一,然而到目前为止,它可能仍然是在实际运用中如何进行网络分析的最超前的例子。莱文的研究把自我中心网和社会中心网关注的问题整合在一个具有凝聚力的模型中,即把多个网络嵌入到一个多维空间模型之中。它开拓了新的研究方向,提供了一种研究框架,在该框架中,图论的一些重要概念——密度、中心性和所有类型的社会圈——可与用来描述区域性的和工业部分的一些位置概念结合在一起,并且可以在复杂但却易懂的图中得到体现。在本书中,笔者的目的是澄清一些重要理论观念,这些理论观念必须以这种方法论和理论综合在一起的方式描述出来。

注 释

1 在附录中描述的 KRACKPLOT 程序包含了一个用来建构这些环形图的程序,尽管排列这些点的顺序仍要由研究者来决定。

2 多维量表也可用来揭示属性数据的特征,但这超出了本书的范围。

3 欧几里德量纲还有额外的功能,此处无需过多讨论。总的来说,一个欧几里德量纲允许使用所有常见的算术运算(加、减、乘、除),并且它也符合传统的学院几何学(例如,毕达哥拉斯定理)原理。可以构建一些量纲模型(而不是欧几里德模型),但它们在社会网络分析中还不是特别重要。一个这样的模型是曲线空间"黎曼式的"量纲(黎曼(Georg Friedrich Bernhard Reimann,1826-1866),德国数学家。1954 年演说的《关于几何基础的假设》建立了黎曼空间的概念,后来成为广义相对论的数学基础。我们的宇宙空间就是黎曼空间,光线在黎曼空间中是弯曲的——译注)(Reimannian metric),毕

达哥拉斯定理以及其他我们熟悉的定理在该空间中是不适用的。

4　这一点描述了量纲式多维量表的特征。与之稍有不同的非量纲式多维量表方法将在本章后文讨论。从大多数实践的目的来讲，一种似乎能产生与 MDS 类似结果的程序是弹性嵌入（spring embedding）程序（Kamada and Kawai，1989）。虽然它在 PAJEK 程序中使用过（参见附录），但是其基础还没有在二手文献中谈及。

5　有时“不相似性”又被称为“差异性”。

6　就 A 的位置来说，确实恰恰存在一个类似的反射问题。如果初始线被画成 BA 方向而不是 AB 方向（即沿着相反的方向），那么线 ABC 将会沿着线 BC 反射。感兴趣的读者可以亲自尝试。

7　有一点很重要，即不要把在主成分分析中使用的“成分”概念与图论中的“成分”概念相混淆。在图论中，一个成分是图中的一个特定部分；在主成分分析中，一个成分指的是“维度”或“因素”。

8　重要的是不要把 Q-模型分析与阿特金的 Q-分析相混淆，后者在第六章中有所涉及，本章稍后也会提到。

9　对于一个不对称的邻接矩阵来说，常常可以产生一个联合的行和列的解，该解可以表示由一些案例占用的联合空间。

10　参见考克森（Coxon，1982）、克鲁斯卡尔和威什（Kruskal and Wish，1978）。关于非量纲式多维量表的初始文献有谢泼德（Shepard，1962），克鲁斯卡尔（Kruskal，1964a，b），格特曼（Guttman，1968）及林格斯和洛斯坎（Lingoes and Roskam，1973）。

11　由于最初的型构只是为检验而生成的，所以其来源并不重要。因此，一个量纲程序可用来生成这种初始型构。

12　严格地讲，这已经过度简化了，但它与检验拟合优度时使用的一般原理相对应。最一般的原则是数据秩（等级）和距离秩（等级）之间的关系是“单调的”——也就是说，即使不一定与轴形成 45 度角，线还是单调地向右上方移动。

13　UCINET 和 KRACKPLOT 能够构建简单的环形社群图，但克雷姆珀尔自己的算法现在还没有得到普遍应用。

14　UCINET 带有一些换算程序，该程序用来将它自己的输出数据文件转换为 MAGE 和许多其他分子建模程序所要求的格式。

15　该矩阵虽然是方阵，但却是不对称的，因为列代表各个回答者，而行代表的是这些回答者宣称的与自己互动之人。在此意义上，数据是从列指向行的。因此，劳曼使用的是格特曼-林格斯（Guttman-Lingoes）程序的非对称变体，该版本通常用于长方形的发生阵之中。

16　注意，试图把球形域的中心等同于第 5 章讨论的任何网络中心，这从根本上说是错误的。事实上，在莱文的分析中的“中心”与劳曼和帕佩的著作讨论的“中心”恰恰相反。

附录 社会网络软件包

本书的许多地方都提到，在社会网络分析软件中，有不同的程序可资利用。本附录的目的就是介绍一些研究者易于获得的主要软件，并试着说明它们的主要特性。从使用者的角度看，笔者将讨论可在个人电脑上使用的如下四种主要程序：GRADAP、STRUCTURE、UCINET 和 PAJEK。

【GRADAP】 GRADAP 即"图的定义与分析程序（Graph Definition and Analysis Package）"。它最初是在 CDC 公司的 Cyber 系列的主机型电脑（CDC Cyber mainframe computers）上使用①的，而其微型机版本则始于 1988 年。该程序已经做了少许的改进，但基本上同样的版本还可以使用[1]。这个程序可以在一个最低配置的电脑上运行，但需要有一个数学协处理器（这种处理器在大多数 486 和所有的奔腾电脑上都不可或缺）。该程序要求的内存极小，因为它要在硬盘上建立一个覆盖（overlay）文件。它可以直接在 DOS 中运行，或者在 Windows 的任何版本中的 DOS 下运行。它的使用手册内容全面，有 580 页，但没有用简单的术语告诉使用者怎样实际操作。

GRADAP 不是交互式的，而是以批处理模式（batch mode）运行的。它的设计兼容了 SPSS 数据文件和命令语言，运行方式与 SPSS 极其相似，有一个输入的"程序文件"和一些独立的数据文件。各种指令和数据要从输入文件中读取，而如果一切顺利的话，使用者就在被通知程序已经执行完毕之前，会看到开启的屏幕和大量的硬盘活动。（令人相当不解的是，开启的屏幕说它"正在装载程序"，即使分析正在运行的时候也如此。）输入命令文件包含一些有关运行这个程序的指令。各种命令的书写格式很接近于 SPSS 的语言和句法。命令文件把"程序（PROCEDURE）"、"选项（OPTION）"和"统计（STATISTICS）"栏分开，以便具体指定到底要执行什么程序。

分析的结果写在一个新的输出文件中。这是一个常规文本文件，可以打印出来或在屏幕上观看。GRADAP 是以 132 列宽的形式操作的，而不是 80 列的

---

① 1959 年，美国明尼苏达州的 CDC 公司（Control Data Corporation，直译是"控制信息公司"）制造了世界上第一台商业主机型电脑（mainframe computer），主机型电脑是一种控制信息的"中心化机构（central institution）"，是政府、企业垄断性处理信息的设施——译注。

屏幕格式，因此如果希望输出文件可读的话，就有必要用较小的字体（不成比例地）打印。输出文件包含了在输入命令文件中要求的全部计算和列表。

GRADAP 分析所要求的基本数据是由一个点集和一个线集组成的。点集包含网络中列举出来的各个点，以及有关它们的特征方面的信息（记作 pointinfos）。“Pointinfos”的例子包括各个点所代表的人名，他们在回答问卷的各项问题或访谈提纲时所作的回答。线集列举出由两个点连接的各条线，它通过指定一点为“头”，一点为“尾”，从而允许线上带有方向。与在点集中一样，“lineinfos”选项也可以加在每一条线上，以此来表现出诸如关系的类型、位置及持续时间等特性。有关 GRADAP 数据文件格式的进一步信息，可在本书第 3 章中找到。

为使任何分析成为可能，必须根据点和线数据来建构图。使用者为图起一个名，并用该名字来指称它。在界面背后，GRADAP 把初始数据转变为内部系统文件——每个图被转换为 5 个文件。它们被赋予的 DOS 扩展名为“. GSF”，但是文件一旦被创建，它们一般就与使用者无关了[2]。在创建了一个图之后，就可以利用一个输入命令文件中的“USE GRAPH *name*”这个命令对一些同样的图系统文件进行深入分析，这里的“*name*”是为图起的名称。利用这个命令，GRADAP 会自动导出系统文件并运行它们。可对系统文件进行备份，或者利用“SAVE GRAPH *name*”这个命令保存下来。该命令把这些系统文件写成一个大文件。用“GET GRAPH *name*”[3]这个命令就可以从这个文件中再构建一些可运行的系统文件。

在 GRADAP 中执行社会网络分析的主要普遍目标程序包括“子图（SUBGRAPH）”，“中心度（CENTRALITY）”，“邻接性（ADJACENCY）”和“距离（DISTANCE）”。在具体指定图和任务之后，必须在一个输入命令文件中命名所选择的程序，或者立刻紧随着另一个已完成的程序。“程序（PROCEDURE）”这一栏具体指定一些特定的分析过程，其后是“选择（CHOICE）”栏，并且在某些情况下随后跟着一个“统计（STATICTICS）”栏，该栏确切地指定了如何补充这个特定的程序。

“子图（SUBGRAPH）”选项由一系列用来确定无向图和有向图中的派系、成分和块（即环成分）的程序组成，使用者可以设定各种最小规模和距离。所提供的各种选项可以具体地指定输出结果的详细程度和列举存放地址。对于区分出来的每一个子图来说，输出信息一般包括它的规模和密度，还有它的成员名单（如果选择了这项的话）。

中心度（CENTRALITY）选项是用来计算各个点的局部中心度和整体中心度，以及整个网络的密度和中心势（centralization）。在该项中可以选择各种点的中心度和图的中心势指数进行计算。用“统计（STATISTICS）”一栏可以探索诸如分散值（spread）和衰减值（attenuation）这样的主题。

用“邻接性（ADJACENCY）”一栏可构建邻接矩阵，考虑到线的多样性，所用的选项可以指明矩阵本身是否根据每个点的邻点来具体详细地列举出来，如果在有向图中，也要详细指明是否根据各个点的点入度和点出度来列举。用来分析这个程序的“统计（STATISTICS）”一栏允许使用者详细指定在输出结果中各个点是如何分类的（例如是按照名称，还是按照参考数）。

最后，我们可用"距离（DISTANCE）"栏来构建距离矩阵，可以导出诸如各个点的"可达性（reachability）"，以及这些点与其他点之间的距离的总和、中位值和平均值等。"统计（STATISTICS）"一栏决定了如何整理矩阵以及是否计算出作为一个整体的网络的直径和其他参数。

邻接矩阵和距离矩阵可以打印出来，成为可用其他软件（如与多维量表和聚类分析等有关的软件）分析的文件。这些程序本身在 GRADAP 程序中是不提供的。然而，这个程序兼容了 dBase 和 INGRES 的数据库格式。

总的来说，GRADAP 的效率较低，也不太容易使用。它有着作为开创性软件的所有缺点，并且不幸的是，它的开发者没有给出更新的版本，然而，该程序总体上根植于图论，提供了大量可供分析的测度，输出结果比较直观，因而比较容易理解。

【STRUCTURE】 STRUCTURE 开创于 1975 年，是罗纳德·伯特（Ronald Burt）为他自己有关结构自主性（structural autonomy）的研究而开发的。从此之后，他进一步扩充，使之成为一套功能强大的程序[4]。到目前为止，它有两种不同的版本：一个是基础版（Basic Edition），该版本几乎可以在所有微机中运行；一个是虚拟存储版（Virtual Memory Edition），该版本需要在有 1 兆的随机存取存储器和 5 兆的可用硬盘空间的 386 或更好的机型中使用。基础版在可加以分析的网络规模上有限制（最多 85 个行动者），因为它只在 RAM 中才可以运行；而虚拟存储版则可以在硬盘中来回读取数据，因而可以处理更大的网络（硬盘中需要有 53 Mb 的可用空间，可处理 999 个行动者的数据）。近年来，随着硬盘容量的不断增加，越来越多的使用者可用虚拟存储版来分析网络。在线手册和打印版手册都可得到。在准备命令文件的时候，有一个"辅助（ASSISTANT）"程序，而在处理数据文件的时候，有"编辑（JEDIT）"程序可用。

与 GRADAP 一样，STRUCTURE 也是以批处理模式（而非交互式）运作的，所以有必要建立一个输入命令文件、一个数据文件和一个输出文件。在 DOS 提示界面（该程序没有 Windows 版本）下输入"STRUC"，就会装载基础程序，然后在提示下键入"输入命令文件"名。这个文件会告诉 STRUCTURE 用哪个数据文件，进行什么样的特定分析，以及把分析结果存储起来的输出文件名是什么。为了运行虚拟存储版，需要键入"VSTRUC"，这将展示给使用者一个菜单系统，该菜单系统的使用有一定的交互性。

命令文件只是一个文本文件，因而可以在任何文本编辑器中创建，只要该编辑器有一个文字处理程序，或者有一个"辅助（ASSISTANT）"程序就行。该文件包含大量独立的命令栏，"ASSISTANT"项有助于编辑这些命令（它也内建了一些用于检验对等性假设的程序，可进行蒙特卡罗［Monte Carlo①］）分析。在一

---

① 蒙特卡罗（Monte Carlo）法是用来解决数理问题的非确定性的（概率统计的或随机的）数值方法——译注。

个命令文件中，核心栏包含“数据（DATA）”，“网络（NETWORK）”和“分析（ANALYZE）”命令。这些命令对数据文件进行命名，描述数据，确定在分析中所使用的一些特定程序。因此，就最简单的形式而言，一个命令文件只包含这三条命令。通过构建这三栏命令序列，就可以进行更复杂的分析，每个ANALYZE命令都指令程序按要求进行分析。然而，一个针对自身的ANALYZE命令仅仅执行一些默认的计算，该栏一般根据一些特定的命令和选项进行分析。主要的程序是派系（CLIQUES）、权力（POWER）、位置（POSITIONS）、自主性（AUTONOMY），以及那些用来计算密度和其他图参数的选项。数据文件只是一些简单的矩阵，可以像创建命令文件一样创建一个数据文件，也可以在那种能够写输出矩阵的数据表或数据库程序中创建数据文件。也可能用GRADAP构造一些可直接输入到“STRUCURE”之中的矩阵。然而，也可以选择JEDIT程序来完成这件事。

该程序本身是在伯特关于凝聚力（cohesion）、结构对等性（structural equivalence）、声望（prominence）、范围（range）和经纪人（brokerage）等观点的基础上构建起来的。这意味着它和作为GRADAP和UCINET基础的一些图论假设有所不同。例如，POSITIONS命令被用来执行结构对等性分析，而AUTONOMY命令则用来分析结构独立性和限制性①（constraint）[5]。该程序还能以一种查错的模式运行，以便检查命令语句和数据文件的格式。这就避免了程序进行长时间的运作，也不会因为命令文件的一个句法错误而导致全盘崩溃。笔者并没有对该程序在不同网络中的运行速度进行全面检验，但是对于小型网络来说，基础版要比虚拟存储版更快些。这个差别在很大程度上是因为设置虚拟存储版文件的时候要花一些时间。因此，在分析小型数据的时候，使用虚拟存储版可能会稍微多用一些时间，但是在分析大型数据时其作用是不可忽视的。

STRUCTURE软件易用，易维护，但它缺少其他程序中拥有的广泛的图论程序。该软件专攻于位置分析，这个能力弥补了上述缺陷。

【UCINET】　UCINET软件是由加州大学欧文（Irvine）分校的一群网络分析者编写的。现在对该软件进行扩展的团队是由斯蒂芬·博加提（Stephen Borgatti）、马丁·埃弗里特（Martin Everett）和林顿·弗里曼（Linton Freeman）[6]组成的。该软件最初是一组用Basic语言编写的模块，渐渐地发展成综合性的DOS程序，现在已经可作为一种Windows程序来使用了。这是一个具有一个通

---

① 从概念上讲，一个人的“限制性（constraint）”指的是此人在自己的网络中运用结构洞的能力或者协商的能力。“你自己的机会受到的限制取决于，(1)你曾经投入了大量网络时间和精力的另外一个接触者q，在多大程度上向(2)接触者j的关系投入大量的精力。”上述说明可以操作化为$C_{ij}=(p_{ij}+\sum_{q}p_{iq}p_{qj})^2$。其中$p_{iq}$是在行动者$i$的全部关系中，投入到$q$的关系占总关系的比例（Burt，R. S. 1992. *Structural Holes*. Harvard University Press. P. 54）——译注。

用目标、易于使用的程序，它还涵盖了一些基本的图论概念、位置分析法和多维量表分析法等。在我看来，它是当今最好使用的程序，最适合新手使用。这个程序可应用于任何一个现代微机之中，只要它至少有1.5 Mb的 RAM。存储容量非常重要，因为该程序运行在任何适当的虚拟存储器中，放弃追求处理大数据的速度。在一些基本的派系程序中，它最多可处理拥有 500 个点的数据，尽管诸如多维量表这样的程序只能运行在小网络中。

UCINET5.0 版的数据文件和 4.0 版的一样，但是它们与 3.0 版使用的文件稍有区别。这些数据文件以矩阵形式存储，由一些简单的字母数字文件组成。数据文件中的各行代表发生阵或邻接矩阵中的各个行，但是标题行包含了行数、列数以及它们所用的标签的详细信息。该程序还包含一些内建步骤，可用来转换早期的 UCINET 数据文件，它也可以把 STRUCTURE 和 NEGOPY 文件转换成 UCINET 格式的文件。除了可用各种形式进行输出之外，还有大量的转换程序可资使用，这使得 UCINET 几乎完整地进入其他社会网络分析程序之中。

除了可用一系列命令来进行文件管理和设定程序选项之外，菜单栏还有 4 个主要选项："数据(DATA)"，"转换(TRANSFORM)"，"网络(NETWORK)"和"工具(TOOLS)"。DATA 和 TRANSFORM 这两个选项结合在一起，可执行几乎所有的数据管理任务：输入，转换和输出都是这样进行的。

创建数据文件的最简单的办法就是用直觉的和内建的(built-in)数据表格式的数据输入系统，这可以从 DATA 菜单中或者工具栏的一个按钮上获得。这需要一个关联列表格式(linked list format)，即对于每个点来说，它可以显示与该点相连的所有其他点的编码值。除了利用 UCINET 空白表进行输入和编辑之外，也可以从 Excel 工作表中输入(输出)数据。初始数据输入之后，就可以对数据文件进行编辑，可以执行各种重排和转换分析来区分出各个子集合以便于进一步分析。例如，可以对各行、各列进行重排，分类，转置，也可以对各条线的权数加以改变。这最后一步——即所谓对矩阵进行"二值处理(dichotomizing)"——使得一系列数据文件的准备工作更加容易，这些数据文件可用于进行诸如嵌套成分分析等操作。

主要的社会网络分析程序出现在 NETWORK 这个目录下，它的子目录有："凝聚力(COHESION)"分析，"成分(COMPONENTS)"分析，"中心性(CENTRALITY)"分析，"子群(SUBGROUPS)"分析，"角色和位置(ROLES & POSITIONS)"分析等，还有各种更专业的程序。利用 COHESION 可以计算一些基本的有关线的计算，如途径(paths)，距离(distances)和捷径(geodesics)等，一个独立的 PROPERTIES 菜单可以计算密度。CENTRALITY 项可以计算各种点度中心度、接近中心度、中间中心度测度以及其他中心度和声望测度。SUBGROUPS 菜单可以为区分出 n-派系(n-cliques)，n-宗派(n-clans)和 k-丛(k-plexes)提供大量强大的技术，而 COMPONENTS 选项则可以检测到一些简单的成分、环成分和 k-核等。补充这些图论测度的是"角色和位置(ROLES & POSITIONS)"目录下的结构对等性测量。在这个目录下，可以进行 CONCOR 和"规则对等性分析(REGE)"分析，还有其他位置分析算法。最后，用 TOOLS 选

项可进行量纲式的和非量纲式的多维量表分析、聚类分析、因素分析和对应分析。这些程序可满足很多目的的分析。尽管对社群图的恰当的视觉检查意味着把输出信息转变为一种更专业的程序，但是这些程序的输出结果表现在屏幕上就是散点图或树状图。

【PAJEK】　PAJEK——斯洛文尼亚语为蜘蛛之意——是一种专门用来处理超大型数据集合的软件。该软件是由弗拉迪米尔·巴塔格烈（Vladimir Batagelj）和安德里·穆瓦（Andrej Mrvar）编写的，于1996年底开始免费使用，并定期升级。由于该软件一直在发展，所以它仍然没有正式的解密版[7]。

该软件是一个Windows程序，它在一个主窗口和各种子窗口中展示其分析工作和结果。与UCINET软件中的DATA和TRANSFORM这两个选项对应，在PAJEK中存在"文件（FILES）"和"网络（NET）"选项。FILES菜单还有一些选项，如"读取（READ）"，"编辑（EDIT）"或者对数据文件进行"分类（SORT）"等。与UCINET数据文件在格式上类似，它们在形式上可以是邻接矩阵本身，也可以是对数据进行分区或聚类分析的结果。利用来自NET的一些命令，可以对网络进行转置或简化处理。正是在这里隐藏着用来检测成分的命令。其他许多菜单选项也可以进行各种分区和聚类，它们是专门设计用来减少巨型网络的规模的，以使之易于分析。例如，可以对一个巨大的网络进行分析和分区，然后对每个分区单独进行更细致的分析。

PAJEK并不包含UCINET或GRADAP中出现的大量网络测度，但它确实可以对大型网络进行有效的分析。然而，对于许多使用者来说，该程序的最令人感兴趣之处是在DRAW这个菜单下的各种选项。正是在这里，使用者才可以选择一些程序在屏幕上画出二维和三维的社群图。与MDS相似，PAJEK也使用一种弹性（spring）嵌入程序，可以对作为结果的图进行着色处理，确定标签，以便突出它的一些核心特征。还可利用一些选项对社群图进行平面旋转或空间旋转，以便从不同角度对图进行观察，同时用鼠标也可以轻松地移动每个点。所有这些操作都可以进行非常细致的控制。所生成的社群图可以用各种形式导出，包括Postscript形式①（便于打印），分子矩阵（MOL②）形式（便于分子建模）和"虚拟现实建模语言（VRML③）"形式（便于在网上浏览）。

虽然这个程序的说明性文件很有限，但是PAJEK软件对那些分析大型数据（显然，可以对上千个点进行有效的分析）来说非常有用，也适用于那些想寻找

---

① PostScript是一种编程语言，适用于列印图像和文字（无论是在纸，胶片或非物质的CRT都可）。它是页面描述语言，于1985年由Adobe推出，其最大特点是避免了版面制作对使用设备的依赖性，能够综合处理文字和图形、图像，因而对印刷业产生了深远影响——译注。

② MOL应该为分子矩阵（Molecular Matrix）格式的缩写——译注。

③ VRML应为Virtual Reality Modeling Language的缩写——译注。

一个简便可行的方法使网络可视化的人。这个程序在将来有巨大的发展潜力。它的一个显著特征就是宏工具(macro facility),这是对 GRADAP 和 UCINET 中命令文件的一个改进,它可以记录各种操作并加以保存,以便后续运用。这给其他程序指明一个方向,即逐渐建立至少一条基础性的宏语言,以便控制程序操作的方式。

【其他软件】

KRACKPLOT:该软件由大卫·克拉克哈特(David Krackhardt)、凯瑟琳·麦格拉斯(Cathleen McGrath)和吉姆·布莱斯(Jim Blythe)编写,并由分析技术公司(Analytical Technologies)[8]发布。其数据输入文件可以由 UCINET 文件转换而来,或者直接来自于文本编辑器。它是专门设计用来在屏幕上绘制数据图的软件,可以形成圆形图和"多维量表"(MDS)图。可以对点作标记,也可以用鼠标移动点。可以显示、打印社群图,或者以 GIF 格式保存社群图。

NETIMAGE:该软件由林顿·弗里曼(Linton Freeman)[9]创建。它是一个小程序,用来制造图的照片制版像(camera-ready images of graphs)。它有一系列程序可以在 Windows,UNIX 或 LINUX 系统中运行。当运行主要程序时,一个窗口就显示出社群图。输入的数据可以是 UCINET 的输出文件或简单文本文件。

NEGOPY:由比尔·理查德(Bill Richards)创建[10],是最早的网络分析程序之一。它可以在 DOS 或 Windows 界面中运行,最多可以处理1 000个点或 20 000 条线。它专长于检验那些既在派系上,也在位置上具有相似性的子群体。另一个来源相同的相关程序(MULTINET)可进行各种形式的结构对等性分析,并且是专门处理个体中心网络数据的。现在,一种允许输入 UCINET 和 KRACKPLOT 文件的版本正在酝酿之中。

使用者最终主要选用什么样的程序,这是一个个人偏好问题,或许也是个人的资金筹措问题。GRADAP 似乎更易于在电脑上处理大型的数据,它是围绕我们熟悉的一些图论原则建立起来的。STRUCTURE 软件更易于使用,这要感谢"辅助(ASSISTANT)"程序,它可以用蒙特卡罗法进行模拟测试,只是它的一般目标程序比较有限。PAJEK 还在更新发展着,但可以处理超大型的数据,使用起来也很简单。UCINET 有许多有力的特征,其当下的版本可在 Windows 界面下交互式地执行,它提供大量测度,计算速度快并且效率高。

新软件每时每刻都在产生,并且经常以创新性的——有时候是我们不熟悉的方法和测度为基础。只要你很清楚它们要干什么,就很值得逐一对它们进行检测。大部分程序都可以在 INSNA 主页及其链接上看到[11]。许多新的程序也会通过 SOCNET 信息服务网站公布出来[12]。

注 释

1 GRADAP,2.10 解密版(1992 年 2 月 10 日发布)。可从荷兰格罗宁根大学的如下地址获得:iec ProGAMMA,PO Box 841,9700AV,Grőningen,The Netherlands;传真:+31-50-363687;电子邮件:gamma.post@gamma.rug.nl。磁盘和手册的价格为 795 美元。有关 GRADAP 的进一步信息可以通过互联网找到,网址是http://www.gamma.rug.nl/files/p315.html①。

2 GRADAP 命令用"图"这个一般术语来指称与某个特定数据集合相关的全部系列的文件。这使得它们都可以通过单个名字被提及。然而,这一术语在图论中超出了其技术上的含义,在图论中它指的是一个网络的数学模型。GRADAP 的概念最好被称作"图文件"。

3 "UNLOAD"可以用来代替"SAVE","RELOAD"用来代替"GET"。

4 STRUCTURE,基础版,解密版 4.2。可从哥伦比亚大学的如下地址获得:Center for the Social Sciences,420 West 118th street,8th floor, Columbia University,NY 10027,USA。基础版可以通过如下地址下载:http://www.columbia.edu/cu/css/download.htm②,在线使用手册可以从伯特在芝加哥大学的网址上下载,地址为http://www.uchicago.edu/fac/ronald.burt/teaching/strucmanual.pdf③。

5 这些观点在伯特(Burt,1982)中有讨论。

6 UCINET5.0 版本(1999 年 2 月发布)。可从如下地址获得:Analytical Technology,11 Ohlin Lane,Harvard,MA 01451,USA④;传真:+978-456-7373。价格为 150 美元(从 4.0 版本升级价为 100 美元;学生价为 40 美元)。UCINET 的网址是http://eclectic.ss.uci.edu/~lin/ucinet.html⑤。

7 PAJEK 可以从如下网址免费下载:http://vlado.mat.uni-lj.si/vlado/vlado.htm⑥。还可以下载简要的手册。巴塔格烈(Batagelj)的联系方式为:Department of Mathematics, University of Ljubljana,Jadranska 19,1000 Ljubljana,Slovenia;传真:+386-66-217-281;e-mail:vladimir.batagelj@uni-lj.si。

---

① 该网址将自动转到 http://www.scienceplus.nl/scienceplus/main/softwareshop/products.jsp,此新网址销售超过 30 种科学研究方面的软件,其中 GRADAP 已经更名,可能变为 Grapher,关于 GRADAP 的电子邮件地址也有所变动,参见本译注中给出的网站——译注。

② 该网址已经变为 http://faculty.chicagogsb.edu/ronald.burt/teaching/STRUC.EXE——译注。

③ 该网址已经变为 http://faculty.chicagogsb.edu/ronald.burt/teaching/STRUCmanual.pdf——译注。

④ 该地址现已改为 PO Box 920089, Needham, MA 02492 USA——译注。

⑤ 该网址已经迁到http://www.analytictech.com/downloaduc6.htm,原网址已不再使用——译注。

⑥ 该网址现在已变为http://vlado.fmf.uni-lj.si/vlado/vlado.htm。巴塔格烈的地址和电子邮件也变为 vladimir.batagelj@fmf.uni-lj.si。实际上,在 UCINET 中已经挂有 PAJEK 软件——译注。

8 其地址参见上述注释6。3.0版本的价格为125美元(学生价为39美元)。相关信息可以从http://www.heinz.cmu.edu/~krack 上获得①。

9 正式的名字用到了大写和小写字母,并且拼写为NetImage这个名字。在网站http://tarski.ss.uci.edu/netim.html 上可以免费获得②。

10 4.302版本(1995年11月发布)。一个小型演示片和详尽的手册可以免费下载,完整的程序可向http://www.sfu.ca/~richards/negopy.htm③ 订购。

11 INSNA的主页地址为http://www.heinz.cmu.edu/project/INSNA④。

12 如订购这一服务,可以发送电子邮件到 listserv@nervm.nerdc.ufl.edu,邮件内容可以简单地写成"订阅 SOCNET 你的名字",用你自己的名字替代"你的名字"即可。

---

① 该地址已变为http://www.andrew.cmu.edu/user/krack——译注。

② 该网址已经不可用,但是在http://www.insna.org/INSNA/soft_inf.html 这个网址上可以找到很多与之相近的软件——译注。

③ 该地址不再使用,改为http://www.sfu.ca/%7Erichards/Pages/negord.htm——译注。

④ 该网址已变为http://www.insna.org——译注。

参考文献

Abell, P. (1986) The *Syntax of Social Life*. Oxford: Oxford University Press.

Alba, R. D. (1973) 'A Graph Theoretic Definition of a Sociometric Clique', *Journal of Mathematical Sociology*, 3.

Alba, R. D. (1982) 'Taking Stock of Network Analysis: A Decade's Results', *Research in the Sociology of Organizations*, 1.

Alba, R. D. and Kadushin, C. (1976) 'The Intersection of Social Circles: A New Measure of Social Proximity in Networks', *Sociological Methods and Research*, 5.

Alba, R. D. and Moore, G. (1978) 'Elite Social Circles', *Sociological Methods and Research*, 7.

Allen, M. P. (1980) 'Cliques Versus Clusters in Corporate Networks', Paper presented to the Pacific Sociological Association, San Francisco.

Anderberg, M. R. (1973) *Cluster Analysis for Applications*. New York: Academic Press.

Anthonisse, J. (1971) *The Rush in a Directed Graph*. Amsterdam: University of Amsterdam Mathematical Centre.

Arabie, P., Boorman, S. A. and Levitt, P. R. (1978) 'Constructing Blockmodels: How and Why', *Journal of Mathematical Psychology*, 17.

Arensberg, C. M. and Kimball, S. T. (1940) *Family and Community in Ireland*. London: Peter Smith.

Aron, R. (1964) *German Sociology*. Glencoe: Free Press.

Atkin, R. (1974) *Mathematical Structure in Human Affairs*. London: Heinemann.

Atkin, R. (1977) *Combinatorial Connectivities in Social Systems*. Basle: Birkhauser.

Atkin, R. (1981) *Multidimensional Man*. Harmondsworth: Penguin.

Bailey, F. G. (1969) *Stratagems and Spoils*. Oxford: Basil Blackwell.

Bailey, K. D. (1976) 'Cluster Analysis', in Heise (ed.), 1976.

Bales, R. F. (1950) *Interaction Process Analysis*. Reading, MA: Addison-Wesley.

Banck, G. A. (1973) 'Network Analysis and Social Theory', in Boissevain and Mitchell (eds), 1973.

Barnes, J. A. (1954) 'Class and Committee in a Norwegian Island Parish', *Human Relations*, 7.

Barnes, J. A. (1969) 'Graph Theory and Social Networks', *Sociology*, 3.

Barnes, J. A. (1974) *Social Networks*. Module in Anthropology, No. 26. Reading, MA: Addison-Wesley.

Barnes, J. A. and Harary, F. (1983) 'Graph Theory in Network Analysis', *Social Networks*, 5.

Bavelas, A. (1948) 'A Mathematical Model for Group Structure', *Applied Anthropology*, 7.

Bavelas, A. (1950) 'Communication Patterns in Task-Oriented Groups', *Journal of the Acoustical Society of America*, 22.

Bearden, J., Atwood, W., Freitag, P., Hendricks, C., Mintz, B. and Schwartz, M. (1975) 'The Nature and Extent of Bank Centrality in Corporate Networks'. Paper to the American Sociological Association.

Beauchamp, M. (1965) 'An Improved Index of Centrality', *Behavioral Science*, 10.

Beaumont, J. C. and Gattrell, A. C. (1982) *An Introduction to Q-Analysis*. Norwich: Geo Publications.

Berkowitz, S. D. (1982) *An Introduction to Structural Analysis*. Toronto: Butter-worths.

Berkowitz, S. D. and Heil, G. (1980) 'Dualities in Methods of Social Network Analysis'. Discussion Paper, University of Toronto Structural Analysis Program.

Berkowitz, S. D., Carrington, P. J., Corman, J. S. and Waverman, L. (1979) 'Flexible Description for a Large Scale Corporate Data Base', *Social Networks*, 2.

Blau, P. M. (1977a) *Inequality and Heterogeneity*. New York: Free Press.

Blau, P. M. (1977b) 'A Macrosociological Theory of Social Structure', *American Journal of Sociology*, 83.

Blau, P. M. and Duncan, O. D. (1967) *The American Occupational Structure*. New York: John Wiley.

Bogardus, E. S. (1959) *Social Distance*. Yellow Springs, OH: Antioch Press.

Boissevain, J. (1974) *Friends of Friends*. Oxford: Basil Blackwell.

Boissevain, J. and Mitchell, J. D. (eds) (1973) *Network Analysis*. The Hague: Mouton.

Bonacich, P. (1972) 'Technique for Analysing Overlapping Membership', in Costner (ed.), 1972.

Bonacich, P. (1987) 'Power and Centrality: A Family of Measures', *American Sociological Review*, 52.

Bonacich, P. and McConaghy, M. (1979) 'The Algebra of Blockmodelling', in Schuessler (ed.), 1979.

Boorman, S. A. and White, H. C. (1976) 'Social Structure from Multiple Networks, II', *American Journal of Sociology*, 81.

Borgatti, S. and Everett, M. G. (1989) 'The Class of All Regular Equivalences: Algebraic Structure and Computation', *Social Networks*, 11.

Borgatti, S., Boyd, J. and Everett, M. G. (1989) 'Iterated Roles: Mathematics and Applications', *Social Networks*, 11.

Bott, E. (1955) 'Urban Families: Conjugal Roles and Social Networks', *Human Relations*, 8.

Bott, E. (1956) 'Urban Families: The Norms of Conjugal Roles', *Human Relations*, 9.

Bott, E. (1957) *Family and Social Network*. London: Tavistock.

Boyd, J. P. (1969) 'The Algebra of Group Kinship', *Journal of Mathematical Sociology*, 6.

Brass, D. J. and Burckhardt, M. E. (1992) 'Centrality and Power in Organizations', in Nohira and Eccles (eds), 1992.

Breiger, R. L. (1979) 'Toward an Operational Theory of Community Elite Structure', *Quality and Quantity*, 13.

Breiger, R. L. (1981) 'The Social Class Structure of Occupational Mobility', *American Journal*

of Sociology, 87.

Breiger, R. L. (1982) 'A Structural Analysis of Occupational Mobility', in Marsden and Lin (eds), 1982.

Breiger, R. L., Boorman, S. A. and Arabie, P. (1975) 'An Algorithm for Blocking Relational Data, with Applications to Social Network Analysis', *Journal of Mathematical Psychology*, 12.

Brent, E. E. (1985) 'Relational Database Structures and Concept Formation in the Social Sciences', *Computers and the Social Sciences*, 1.

Brint, S. (1992) 'Hidden Meanings: Cultural Content and Context in Harrison White's Structural Sociology', *Sociological Theory*, 10.

Buckley, W. (1967) *Sociology and Modern Systems Theory*. Englewood Cliffs. NJ: Prentice-Hall.

Bulmer, M. (1985) 'The Rejuvenation of Community Studies? Neighbours, Networks, and Policy', *Sociological Review*, 33.

Burt, R. S. (1976) 'Positions in Social Networks', *Social Forces*, 55.

Burt, R. S. (1977a) 'Positions in Multiple Network Systems, Part One', *Social Forces*, 56.

Burt, R. S. (1977b) 'Positions in Multiple Network Systems, Part Two', *Social Forces*, 56.

Burt, R. S. (1979) 'A Structural Theory of Interlocking Corporate Directorates', *Social Networks*, 1.

Burt, R. S. (1980) 'Models of Network Structure', *Annual Review of Sociology*, 6.

Burt, R. S. (1982) *Towards a Structural Theory of Action*. New York: Academic Press.

Burt, R. S. (1983a) 'Studying Status/Role-Sets Using Mass Surveys', in Burt and Minor (eds), 1983.

Burt, R. S. (1983b) *Corporate Profits and Cooptation*. New York: Academic Press.

Burt. R. S. (1983c) 'Firms, Directors and Time in the Directorate Tie Network', *Social Networks*, 5.

Burt, R. S. (1987) 'Social Contagion and Innovation: Cohesion Versus Structural Equivalence', *American Journal of Sociology*, 92.

Burt, R. S. (1988) 'The Stability of American Markets', *American Journal of Sociology*, 94.

Burt, R. S. (1992) *Structural Holes*. New York: Cambridge University Press.

Burt, R. S. and Carlton, D. S. (1989) 'Another Look at the Network Boundaries of American Markets', *American Journal of Sociology*, 95.

Burt, R. S. and Minor, M. J. (eds) (1983) *Applied Network Analysis*. Beverly Hills, CA: Sage Publications.

Burt, R. and Schott, T. (1990) *STRUCTURE4 .1 Reference Manual*. New York: Columbia University Press.

Burt, R. S., Christman, K. P. and Kiburn, H. C. (1980) 'Testing a Structural Theory of Corporate Cooptation: Interlocking Directorate Ties as a Strategy for Avoiding Market Constraints on Profits', *American Sociological Review*, 45.

Carrington, P. J. and Heil, G. H. (1981) 'COBLOC: A Hierarchical Method for Blocking Network Data', *Journal of Mathematical Sociology*, 8.

Carrington, P. J., Heil, G. H. and Berkowitz, S. D. (1980) 'A Goodness-of-Fit Index for Blockmodels', *Social Networks*, 2.

Cartwright, D. and Harary, F. (1956) 'Structural Balance: a Generalisation of Heider's Theory',

Psychological Review, 63. Reprinted in Leinhardt (ed.), 1984.

Cartwright, D. and Zander, A. (eds) (1953) *Group Dynamics*. London: Tavistock.

Christofides, N. (1975) *Graph Theory: An Algorithmic Approach*. New York: Academic Press.

Coleman, J. S., Katz, E. and Menzel, H. (1996) *Medical Innovation: a Diffusion Study*. Indianapolis: Bobbs-Merrill.

Cook, K. S. (1977) 'Exchange and Power in Networks of Interorganisational Relations', *Sociological Quarterly*, 18.

Cook, K. S. (1982) 'Network Structure form an Exchange Perspective', in Marsden and Lin (eds), 1982.

Cook, K. S. and Whitmeyer, J. M. (1992) 'Two Approaches to Social Structure: Exchange Theory and Network Analysis', *Annual Review of Sociology*, 18.

Costner, H. (ed.) (1972) *Sociological Methodology, 1973*. San Francisco: Jossey Bass.

Coxon, A. P. M. (1982) *The Users Guide to Multidimensional Scaling*. London: Heinemann.

Crane, D. (1972) *Invisible Colleges*. Chicago: University of Chicago Press.

Dahl, R. A. (1961) *Who Governs?* New Haven, CT.: Yale University Press.

Daultrey, S. (1976) *Principal Components Analysis*. Norwich: Geo Abstracts.

Davis, A., Gardner, B. B. and Gardner, M. R. (1941) *Deep South*. Chicago: University of Chicago Press.

Davis, J. A. (1967) 'Clustering and Structural Balance in Graphs', *Human Relations*, 20. Reprinted in Leinhardt (ed.), 1984.

Davis, J. A. (1968) 'Structural Balance, Mechanical Solidarity, and Interpersonal Relations', *American Journal of Sociology*, 68.

Doreian, P. (1979) *Mathematics and the Study of Social Relations*. London: Weidenfeld & Nicolson.

Doreian, P. (1980) 'On the Evolution of Group and Network Structure', *Social Networks*, 2.

Doreian, P. (1981) 'Polyhedral Dynamics and Conflict Mobilisation in Social Networks', *Social Networks*, 3.

Doreian, P. (1983) 'Levelling Coalitions in Network Phenomena', *Social Networks*, 4.

Doreian, P. (1987) 'Measuring Regular Equivalence in Symmetrical Structures', *Social Networks*, 9.

Duck, S. and Perlman, D. (eds) (1985) *Personal Relationships*, vol. 1. Beverly Hills, CA: Sage Publications.

Duncan, O. D. (1961) 'A Socio-economic Index for all Occupations', in Reiss(ed), 1961.

Emerson, R. M. (1962) 'Power-Dependence Relations', *American Sociological Review*, 27.

Emerson, R. M. (1964) 'Power-Dependence Relations: Two Experiments', *Sociometry*, 27.

Emirbayer, M. (1997) 'Manifesto for a Relational Sociology', *American Journal of Sociology*, 103.

Emirbayer, M. and Goodwin, J. (1994) 'Network Analysis, Culture, and the Problem of Agency', *American Journal of Sociology*, 99.

Erickson, B. H. (1978) 'Some Problems of Inference from Chain Data', in Schuessler (ed.), 1978.

Erickson, B. H and Nosanchuk, T. (1983) *Understanding Data*. Toronto: McGraw Hill.

Erickson, B. H., Nosanchuk, T. A. and Lee, E. (1981) 'Network Sampling in Practice: Some Second Steps', *Social Networks*, 3.

Erikson, K. T. (1996) *Wayward Puritans.* New York: John Wiley.

Everett, M. G. (1982) 'A Graph Theoretic Blocking Procedure for Social Networks', *Social Networks*, 4.

Everett, M. G. (1983a) 'EBLOC: A Graph Theoretic Blocking Algorithm for Social Networks', *Social Networks*, 5.

Everett, M. G. (1983b) 'An Extension of EBLOC to Valued Graphs', *Social Networks*, 5.

Everett, M. G. (1984) 'An Analysis of Cyclically Dense Data Using EBLOC', *Social Networks*, 6.

Everett, M. G. and Borgatti, S. (1990) 'A Testing Example for Positional Analysis Techniques', *Social Networks*, 12.

Everett, M. G., Boyd, J. P. and Borgatti, S. (1990) 'Ego-Centered and Local Roles: A Graph Theoretic Approach', *Journal of Mathematical Sociology*, 15.

Everitt, B. (1974) *Cluster Analysis.* London: Heinemann.

Fararo, T. J. and Sunshine, M. H. (1964) *A Study of a Biased Friendship Net.* Syracuse, NY: Syracuse University Press.

Faust, K. and Romney, A. K. (1985) 'Does "Structure" Find Structure? A Critique of Burt's Use of Distance as a Measure of Structural Equivalence', *Social Networks*, 7.

Featherman, D. L. and Hauser, R. M. (1978) *Opportunity and Change.* New York: Academic Press.

Festinger, L. (1949) 'The Analysis of Sociograms Using Matrix Algebra', *Human Relations*, 2.

Festinger, L. (1957) *A Theory of Cognitive Dissonance.* Evanston, IL: Row Peterson.

Festinger, L., Riecken, H. W. and Schachter, S. (1959) *When Prophecy Fails.* New York: Harper and Row.

Fischer, C. S. (1977) *Networks and Places: Social Relations in the Urban Setting.* New York: Free Press.

Fischer, C. S. (1982) *To Dwell Among Friends.* Chicago: University of Chicago Press.

Frank, O. (1978a) 'Sampling and Estimation in Large Networks', *Social Networks*, 1.

Frank, O. (1978b) 'Estimation of the Number of Connected Components in a Graph by Using a Sampled Sub-graph', *Scandinavian Journal of Statistics*, 5.

Frank, O. (1979) 'Estimation of Population Totals by Use of Snowball Samples', in Holland and Leinhardt (eds), 1979.

Frank, O. (1988) 'Random Sampling and Social Networks: A Survey of Various Approaches', *Mathematiques, Informatique et Sciences Humaine*, 26.

Frankenberg, R. (1966) *Communities in Britain.* Harmondsworth: Penguin.

Freeman, L. C. (1979) 'Centrality in Social Networks: I. Conceptual Clarification', *Social Networks*, 1.

Freeman, L. C. (1980) 'The Gatekeeper. Pair Dependency and Structural Centrality', *Quality and Quantity*, 14.

Freeman, L. C. (1983) 'Spheres, Cubes, and Boxes: Graph Dimensionality and Network Structure', *Social Networks*, 5.

Freeman, L. C. (1996a) 'Visualizing Social Networks', http://carnap. ss. uci. edu/vis. html

Freeman, L. C. (1996b) 'Using Molecular Modelling Software in Social Network Analysis: A Practicum', http://eclectic. ss. uci. edu/ ~ lin/chem. html

Freeman, L. C., White, D. R. and Romney, A. K. (eds) (1989) *Research Methods in Social Network Analysts.* New Brunswick, NJ: Transaction Books, 1992 edn.

Friedkin, N. E. (1981) 'The Development of Structure in Random Networks', *Social Networks*, 3.

Friedikin, N. (1984) 'Structural Equivalence and Cohesion Explanations of Social Homogeneity', *Sociological Methods and Research*, 12.

Friedkin, N. (1998) *A Structural Theory of Social Influence.* New York: Cambridge University Press.

Galtung, J. (1967) *Theory and Methods of Social Research.* London: George Allen & Unwin.

Gattrell, A. C. (1984a) 'Describing the Structure of a Research Literature: Spatial Diffusion Modelling in Geography', *Environment and Planning*, *B*, 11.

Gattrell, A. C. (1984b) 'The Growth of a Research Speciality', *Annals of the Association of American Geographers*, 74.

Glass, D. V. (1954) *Social Mobility in Britain.* London: Routledge and Kegan Paul.

Goddard, J. and Kirby, A. (1979) *An Introduction to Factor Analysis.* Norwich: Geo Abstracts.

Granovetter, M. (1973) 'The Strength of Weak Ties', *American Journal of Sociology*, 78.

Granovetter, M. (1974) *Getting a Job.* Cambridge, MA: Harvard University Press.

Granovetter, M. (1976) 'Network Sampling: Some First Steps', *American Journal of Sociology*, 81.

Granovetter, M. (1977) Reply to Morgan and Rytina, *American Journal of Sociology*, 82.

Granovetter, M. (1979) 'The Theory-Gap in Social Network Analysis', in Holland and Leinhardt (eds), 1979.

Granovetter, M. (1982) 'The Strength of Weak Ties: A Network Theory Revisited', in Marsden and Lin (eds), 1982.

Grieco, M. (1987) *Keeping It in The Family.* London: Tavistock.

Guttman, L. (1968) 'A General Non-Metric Technique for Finding the Smallest Coordinate Space for a Configuration of Points', *Psychometrika*, 33.

Habermas, J. (1968) 'Labour and Interaction: Remarks on Hegel's Jena Philosophy of Mind', in Habermas, 1974.

Habermas, J. (1974) *Theory and Practice.* London: Heinemann.

Hage, P. and Harary, F. (1983) *Structural Models in Anthropology.* Cambridge: Cambridge University Press.

Hage, P. and Harary, F. (1991) *Exchange in Oceania: A Graph Theoretic Analysis.* Oxford: Oxford University Press.

Hage, P. and Harary, F. (1998) *Island Networks: Communication, Kinship, and Classification Structures in Oceania.* Cambridge: Cambridge University Press.

Harary, F. (1969) *Graph Theory.* Reading, MA: Addison-Wesley.

Harary, F. and Norman, R. Z. (1953) *Graph Theory as a Mathematical Model in Social Science.* Ann Arbor, MI: Institute for Social Research.

Harary, F., Norman, R. Z. and Cartwright, D. (1965) *Structural Models.* New York: John Wiley.

Heider, F. (1946) 'Attitudes and Cognitive Orientation', *Journal of Psychology*, 21. Reprined in Leinhardt (ed.), 1977.

Heise, D. R. (ed.) (1976) *Sociological Methodology,1975* . San Francisco: Jossey Bass.

Heise, D. R. (ed.) (1977) *Sociological Methodology,1978* . San Francisco: Jossey Bass.

Helmers, H. M., Mokken, R. J., Plijter, R. C. and Stockman, F. N. (1975) *Graven Naar Macht.* Amsterdam: Van Gennep.

Holland, P. and Leinhardt, S. (eds) (1979) *Perspecitves on Social Networks.* New York: Academic Press.

Homans, G. C. (1941) *English Villagers of the Thirteenth Century.* Cambridge; MA: Harvard University Press.

Homans, G. C. (1951) *The Human Group.* London: Routledge and Kegan Paul.

Homans, G. C. (1961) *Social Behaviour.* London: Routledge and Kegan Paul.

Honigmann, J. J. (ed.) (1973) *Handbook of Social and Cultural Anthropology.* Chicago: Rand McNally.

Hudson, J. C. (1969) 'Diffusion in a Central Place System', *Geographical Analysis*, 1.

Hunter, F. (1953) *Community Power Structure.* Chapel Hill, NC: University of North Carolina Press.

Johnson, S. C. (1967) 'Hierarchical Clustering Schemes', *Psychometrika*, 32.

Kadushin, C. (1966) 'The Friends and Supporters of Psychotherapy', *American Sociological Review*, 31.

Kadushin, C. (1968) 'Power, Influence and Social Circles: A New Methodology for Studying Opinion Makers', *American Sociological Review*, 33.

Kamada, T. and Kawai, S. (1989) 'An Algorithm for Drawing General Undirected Graphs', *Information Processing Letters*, 31.

Katz, F. (1966) 'Social Participation and Social Structure', *Social Forces*, 45.

Kerr, C. and Fisher, L. H. (1957) 'Plant Sociology: The Elite and the Aborigines', in Komarovsky (ed.), 1957.

Kilworth, P. D. and Bernard, H. R. (1979) 'A Pseudo-model of the Small World Problem', *Social Forces*, 58.

Kim, J. and Mueller, C. W. (1978) *Introduction to Factor Analysis.* Beverly Hills, CA: Sage Publications.

Kline, P. (1994) *An Easy Guide to Factor Analysis.* London: Routledge.

Klovdahl, A. S. (1981) 'A Note on Images of Networks', *Social Networks*, 3.

Klovdahl, A. S. (1986) 'VIEWNET: A New Tool For Network Analysis', *Social Networks*, 8.

Klovdahl, A. S. (1989) 'Urban Social Networks: Some Methodological Problems and Prospects', in Kochen (ed.), 1989.

Knoke, D. and Burt, R. S. (1983) 'Prominence', in Burt and Minor (eds), 1983.

Knoke, D. and Kuklinski, J. H. (1982) *Network Analysis.* Beverly Hills, CA: Sage Publications.

Kochen, M. (ed.) (1989) *The Small World.* Norwood, NJ: Ablex.

Köhler, W. (1925) *The Mentality of Apes*. New York.
Komarovsky, M. (ed.) (1957) *Common Frontiers of the Social Sciences*. Glencoe, IL: Free Press.
König, D. (1936) *Theorie der andlichen und unendlichen Graphen*. New York: Chelsea, 1950 edn.
Krempel, L. (1994) 'Simple Representations of Complex Networks: Strategies for Visualizing Network Structure', http://www.mpi-fg-koeln.mpg.de/~lk/algo5a/algo5a.html
Kruskal, J. B. (1964a) 'Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis', *Psychometrika*, 29.
Kruskal, J. B. (1964b) 'Nonmetric Multidimensional Scaling: A Numerical Method', *Psychometrika*, 29.
Kruskal, J. B. and Wish, M. (1978) *Multidimensional Scaling*. Beverly Hills, CA: Sage Publications
Lankford, P. M. (1974) 'Comparatire Analysis of Clique Identification Methods', *Sociometry*, 37.
Laumann, E. O. (1966) *Prestige and Association in an Urban Community*. Indianapolis: Bobbs-Merrill.
Laumann, E. O. (1973) *Bonds of Pluralism*. New York: John Wiley.
Laumann, E. O. and Pappi, F. U. (1973) 'New Directions in the Study of Community Elites', *American Sociological Review*, 38. Reprinted in Leinhardt (ed.), 1984.
Laumann, E. O. and Pappi, F. U. (1976) *Networks of Collective Action*. New York: Academic Press.
Laumann, E. O., Marsden, P. V. and Prensky, D. (1983) 'The Boundary Specification Problem in Network Analysis', in Burt and Minor (eds), 1983.
Laumann, E. O., Marsden, P. V. and Prensky, D. (1989) 'The Boundary Specification Problem in Network Analysis', in Freeman et al. (eds), 1989.
Layder, D. (1992) *New Strategies in Social Research*. Cambridge: Polity Press.
Lee, N. H. (1969) *The Search for an Abortionist*. Chicago: University of Chicago Press.
Leinhardt, S. (ed.) (1977) *Social Networks: A Developing Paradigm*. New York: Academic Press.
Leinhardt, S. (ed.) (1984) *Sociological Methodology, 1985* . San Francisco: Jossey Bass.
Levine, J. H. (1972) 'The Sphere of Influence', *American Sociological Review*, 37. Reprinted in Scott (ed.), 1990.
Lewin, K. (1936) *Principles of Topological Psychology*. New York: McGraw Hill.
Lewin, K. (1951) *Field Theory in the Social Sciences*. New York: Harper.
Light, J. M. and Mullins, N. C. (1979) 'A Primer on Blockmodelling Procedures', in Holland and Leinhardt (eds) 1979.
Lin, N. (1982) 'Social Resources and Instrumental Action', in Marsden and Lin (eds), 1982.
Lin, N., Dayton, P. N. and Greenwald, P. (1978) 'Analysing the Instrumental Use of Relations in the Context of Social Structure', *Sociological Methods and Research*, 7.
Lingoes, J. C. and Roskam, E. E. (1973) 'A Mathematical and Empirical Analysis of Two Multidimensional Scaling Algorithms', *Psychometrika*, 38.

Lockwood, D. (1956) 'Some Remarks on *The Social System*', *British Journal of Sociology*, 6.

Lorrain, F. and White, H. C. (1971) 'Structural Equivalence of Individuals in Social Networks', *Journal of Mathematical Sociology*, 1.

Luce, R. D. and Perry, A. (1949) 'A Method of Matrix Analysis of Group Structure', *Psychometrika*, 14.

Mariolis, P. (1975) 'Interlocking Directorates and Control of Corporations', *Social Science Quarterly*, 56.

Mariolis, P. and Jones, M. H. (1982) 'Centrality in Corporate Networks: Reliability and Stability', *Administrative Science Quarterly*, 27.

Marsden, P. V. (1982) 'Brokerage Behaviour in Restricted Exchange Networks', in Marsden and Lin (eds), 1982.

Marsden, P. V. and Lin, N. (eds) (1982) *Social Structure and Network Analysis*. Beverly Hills, CA: Sage Publications.

Mayhew, B. H. and Levinger, R. (1976) 'Size and the Density of Interaction in Human Aggregates', *American Journal of Sociology*, 82.

Mayo, E. (1933) *The Human Problems of an Industrial Civilization*. Cambridge, MA: Macmillan.

Mayo, E. (1945) *The Social Problems of an Industrial Civilization*. London: Routledge and Kegan Paul, 1946 edn.

McCallister, L. and Fischer, C. S. (1978) 'A Procedure for Surveying Social Networks', *Sociological Methods and Research*, 7.

McGrath, C., Blythe, J. and Krackhardt, D. (1997) 'Seeing Groups in Graph Layouts', http://www.andrew.cmu.edu/user/cm3t/groups.html

McQuitty, L. (1968) 'Multiple Cluster Types and Dimensions for Interactive Columnar Correlation Analysis', *Multivariate Behavioral Research*, 3.

McQuitty, L. and Clark, J. A. (1968) 'Clusters from Interactive Columnar Correlation Analysis', *Educational and Psychological Measurement*, 28.

Meek, R. L. and Bradley, I. (1986) *Matrices and Society*. Harmondsworth: Penguin.

Milgram, S. (1967) 'The Small World Problem', *Psychology Today*, 1.

Mintz, B. and Schwartz, M. (1985) *The Power Structure of American Business*. Chicago: University of Chicago Press.

Mitchell, J. C. (1969) 'The Concept and Use of Social Networks', In Mitchell (ed.), *Social Networks in Urban Situations*. Manchester: Manchester University Press.

Mizruchi, M. S. (1982) *The American Corporate Network, 1904-1974*. Beverly Hills, CA: Sage Publications.

Mizruchi, M. S. (1992) *The Structure of Corporate Political Action: Interfirm Relationships and their Consequences*. Cambridge, MA: Harvard University Press.

Mizruchi, M. S. (1993) 'Cohesion, Equivalence, and Similarity of Behaviour: A Theoretical and Empirical Assessment', *Social Networks*, 15.

Mizruchi, M. S. (1994) 'Social Network Analysis: Recent Achievements and Current Controversies', *Acta Sociologica*, 37.

Mizruchi, M. S. and Bunting, D. (1981) 'Influence in Corporate Networks: An Examination of Four Measures', *Administrative Science Quarterly*, 26.

Mizruchi, M. S. and Galaskiewicz, J. (1994) 'Networks of Interorganizational relations', in Wasserman ans Galaskiewicz (eds), 1994.

Mizruchi, M. and Schwartz, M. (eds) (1987) *Intercorporate Relations*. Cambridge: Cambridge University Press.

Mokken, R. J. (1974) 'Cliques, Clubs and Clans', *Quality and Quantity*, 13.

Moreno, J. (1934) *Who Shall Survive?* New York: Beacon Press.

Morgan, D. and Rytina, S. (1977) 'Comment on "Network Sampling: Some First Steps" by Mark Granovetter', *American Journal of Sociology*, 83.

Mullins, N. C. (1973) *Theories and Theory Groups in Contemporary American Sociology*. New York: Harper and Row.

Nadel, S. F. (1957) *The Theory of Social Structure*. London: Cohen and West.

Nash, K. and Scott, A. (eds) (1999) *The Blackwell Companion to Political Sociology*. Oxford: Basil Blackwell.

Newcomb, T. (1953) 'An Approach to the Study of Communicative Acts', *Psychological Review*, 60.

Niemeijer, R. (1973) 'Some Applications of the Notion of Density', in Boissevain and Mitchell (eds), 1973.

Nieminen, V. (1974) 'On Centrality in a Graph', *Scandinavian Journal of Psychology*, 15.

Nohira, N. and Eccles, R. G. (eds) (1992) *Networks and Organizations*. Boston, MA: Harvard University Press.

Park, R. E., Burgess, E. W. and McKenzie, R. D. (1925) *The City*. Chicago: University of Chicago Press.

Parsons, T. (1937) *The Structure of Social Action*. New York: McGraw Hill.

Parsons, T. (1951) *The Social System*. Glencoe, IL: Free Press.

Parsons, T., Bales, R. F. and Shils, E. (1953) *Working Papers in the Theory of Action*. Glencoe, IL: Free Press.

Pedersen, P. O. (1970) 'Innovation Diffusion Within and Between National Urban Systems', *Geographical Analysis*, 2.

Price, D. J. de Solla (1965) 'Networks of Scientific Papers', *Science*, 149.

Rapoport, A. (1952) 'Ignition Phenomena in Random Nets', *Bulletin of Mathematical Biophysics*, 14.

Rapoport, A. (1958) 'Nets with Reciprocity Bias', *Bulletin of Mathematical Biophysics*, 20.

Reiss, A. J. (ed.) (1961) *Occupations and Social Status*. New York: Free Press.

Reitz, K. P. and White, D. R. (1989) 'Rethinking the Role Concept: Homomorphisms on Social Networks', In Freeman et al. (eds), 1989.

Roethlisberger, F. J. and Dickson, W. J. (1939) *Management and the Worker*. Cambridge, MA: Harvard University Press.

Roistacher, R. C. (1979) 'Acquisition and Management of Social Network Data', in Holland and Leinhardt (eds), 1979.

Rose, M. (1975) *Industrial Behaviour*. Harmondsworth: Allen Lane.

Rytina, S. (1982) 'Structural Constraints on Interpersonal Contact', in Mardsen and Lin (eds), 1982.

Rytina, S. and Morgan, D. (1982) 'The Arithmetic of Social Relations', *American Journal of Sociology*, 88.

Sabidussi, G. (1966) 'The Centrality Index of a Graph', *Psychometrika*, 31.

Sailer, L. (1978) 'Structural Equivalence: Meaning and Definition', *Social Networks*, 1.

Schuessler, K. F. (1978) *Sociological Methodology,1979* . San Francisco: Jossey Bass.

Schuessler, K. F. (1979) *Sociological Methodology,1980* . San Francisco: Jossey Bass.

Schwartz, J. E. (1977) 'An Examination of CONCOR and Related Methods for Blocking Sociometric Data', in Heise (ed.), 1977.

Scott, J. (1986) *Capitalist Property and Financial Power.* Brighton: Wheatsheaf.

Scott, J. (ed.) (1990) *The Sociology of Elites*, vol. 3. Cheltenham: Edward Elgar.

Scott, J. (1991) 'Networks of Corporate Power', *Annual Review of Sociology*, 17.

Scott, J. (1996) *Stratification and Power: Structures of Class, Status and Command.* Cambridge: Polity Press.

Scott, J. (1997) *Corporate Business and Capitalist Classes.* Oxford: Oxford University Press.

Scott, J. (1999) 'Studying Power', in Nash and Scott (eds), 1999.

Scott, J. and Hughes, M. (1980) *The Anatomy of Scottish Capital.* London: Croom Helm.

Seidman, S. B. (1983) 'Network Structure and Minimum Degree', *Social Networks*, 5.

Seidman, S. B. and Foster. B. L. (1978) 'A Note on the Potential for Genuine Cross-Fertilisation between Anthropology and Mathematics', *Social Networks*, 1.

Sharkey, P. (1989) 'Social Networks and Social Service Workers', *British Journal of Social Work*, 19.

Sharkey, P. (1990) 'Social Networks and Social Service Workers: A Reply to Timms', *British Journal of Social Work*, 20.

Shepard, R. N. (1962) 'The Analysis of Proximities', Parts 1 and 2, *Psychometrika*, 27.

Simmel, G. (1908) *Soziologie.* Berlin: Duncker and Humblot, 1968 edn.

Smith, R. M. (1979) 'Kin and Neighbours in a Thirteenth Century Suffolk Community', *Journal of Family History*, 4.

Snijders, T. A. B. (1981) 'The Degree Variance', *Social Networks*, 3.

Sonquist, J. A. and Koenig, T. (1975) 'Interlocking Directorships in the Top US Corporations', *Insurgent Sociologist*, 5.

Stacey, M. (1969) 'The Myth of Community Studies', *British Journal of Sociology*, 20.

Stokman, F. N., Ziegler, R. and Scott, J. (1985) *Networks of Corporate Power.* Cambridge: Polity Press.

Timms, E. (1990) 'Social Networks and Social Service Workers: A Comment on Sharkey', *British Journal of Social Work*, 20.

Torgerson, W. S. (1952) 'Multidimensional Scaling 1. Theory and Method', *Psychometrika*, 17.

Useem, M. (1984) *The Inner Circle.* New York: Oxford University Press.

Van Poucke, W. (1979) 'Network Constraints on Social Action: Preliminaries for a Network Theory', *Social Networks*, 2.

Warner, W. L. and Lunt, P. S. (1941) *The Social Life of a Modern Community.* New Haven, CT: Yale University Press.

Warner, W. L. and Lunt, P. S. (1942) *The Status System of a Modern Community.* New Haven,

CT: Yale University Press.
Wasserman, S. and Galaskiewicz, J. (eds) (1994) *Advances in Social Network Analysis*. Beverley Hills, CA: Sage Publications.
Weber, M. (1920-21) *Economy and Society*. Berkeley, CA: University of California Press, 1968 edn.
Wellman, B. (1979) 'The Community Question: The Intimate Networks of East Yorkers', *American Journal of Sociology*, 84.
Wellman, B. (1980) 'Network Analysis: From Metaphor and Method to Theory and Substance'. Discussion Paper, University of Toronto Structural Analysis Program. Revised version in Wellman and Berkowitz (eds), 1988.
Wellman, B. (1982) 'Studying Personal Communities', in Marsden and Lin (eds), 1982.
Wellman, B. (1985) 'Domestic Work, Paid Work and Network', in Duck and Perlman (eds), 1985.
Wellman, B. and Berkowitz, S. D. (eds) (1988) *Social Structures*. Cambridge: Cambridge University Press.
Werbner, P. (1990) *The Migration Process*. New York: Bery.
White, D. and Reitz, K. P. (1983) 'Group and Semi-group Homomorphisms on Networks of Relations', *Social Networks*, 5.
White, H. C. (1963) *An Anatomy of Kinship*. Englewood Cliffs, NJ: Prentice-Hall.
White, H. C. (1970) *Chains of Opportunity*. Cambridge, MA: Harvard University Press.
White, H. C. (1992a) *Identity and Control: A Structural Theory of Social Action*. Princeton, NJ: Princeton University Press.
White, H. C. (1992b) 'Social Grammar for Culture: Reply to Steven Brint', *Sociological Theory*, 10.
White, H. C. (1993) *Careers and Creativity: Social Forces in the Arts*. Boulder, CO: Westview Press.
White, H. C., Boorman, S. A. and Breiger, R. L. (1976) 'Social Structure from Multiple Networks, I', *American Journal of Sociology*, 81.
Whitten, N. E. and Wolfe, A. W. (1973) 'Network Analysis', in Honigmann (ed.), 1973.
Willmott, P. (1986) *Social Networks, Informal Care and Public Policy*. London: Policy Studies Institute.
Willmott, P. (1987) *Friendship Networks and Social Support*. London: Policy Studies Institute.
Winship, C. and Mandel, M. (1984) 'Roles and Positions: A Critique and Extension of the Blockmodelling Approach', in Leinhardt (ed.), 1984.
Wu, L. (1984) 'Local Blockmodel Algebra for Analysing Social Networks', in Leinhardt (ed.), 1984.
Yablonsky, L. (1962) *The Violent Gang*. Harmondsworth: Penguin.
Zegers, F. and ten Berghe, J. (1985) 'A Family of Association Coefficients for Metric Scales', *Psychometrika*, 50.

术语（人名）英汉对照表

absolute centre, of a graph （一个图的绝对中心）
absolute density（绝对密度）
adjacency（邻接）
adjacency matrix（邻接矩阵）
Alba, R. D.（阿尔巴）
all attribute cluster（全体属性聚类）
Atkin, R.（阿特金）
attribute data（属性数据）
balance（平衡）
Bales, R. F.（贝尔斯）
bank centrality（银行中心性）
Barnes, J.（巴恩斯）
Bavelas, A.（贝弗拉斯）
Bearden, J.（比尔登）
Berkowitz, S. D.（伯科威茨）
betweenness（中间性）
block（块）
block modeling（块模型）
Bogardus, E.（博格达斯）
Bonacich, P.（伯纳西茨）
Borgatti, S.（博加提）
Bott, E.（博特）
boundaries in networks（网络的边界）
Breiger, R. L.（布雷格）
bridge（桥）
Burt, R.（伯特）
BURT algorithm BURT（算法）
Cartwright, D.（卡特赖特）
centrality（中心性）
centralization（中心势）
Christofides, N.（克利斯托弗兹）
circle（环）
citations in science（科学中的引用）
clan（宗派）
cliques in early Harvard studies（早期哈佛大学研究中的派系）
cliques in graph theory（图论中的派系）
cluster analysis（聚类分析）
community structure（社区结构）
complete linkage cluster（完备关联聚类）
component（成分）;cyclic（环成分）
CONCOR, CONCOR（算法）
configuration in structural analysis（结构分析中的构型）
conflict theory（冲突理论）
confusion over the meaning of（关于块的意义的混淆）
core（核）
core collapse sequence（核塌缩序列）
Crane, D.（克兰）
cut-point（切点）
cycle（环）
Dahl, R.（达尔）
data collection（数据收集）
data matrix（数据矩阵）
Davis, A.（戴维斯）
Davis, J.（戴维斯）
degree（度数）
dentrograms（树形图）

MAGE,(MAGE 画图程序)
Manchester anthropology(曼彻斯特人类学)
Mariolis, P.(马里奥里斯)
matrices(矩阵)
Mayo, E.(梅奥)
McQuitty, L.(麦奎迪)
metric measures(量纲式测度)
Mintz, B.(明茨)
Mitchell, J. C.(米切尔)
Mizruchi, M. S.(米祖齐)
Mokken, R.(莫肯)
Moreno, J.(莫雷诺)
Mullins, N. C.(马林斯)
multidimensional scaling(MDS)(多维量表)
multiplicity(多样性)
n-clan(n-宗派)
n-cliques(n-派系)
Nadel, S. F.(纳德尔)
neighbourhood(邻居关系)
nesting of components(成分的嵌套)
New Haven,(纽黑文市)
Newcomb, T.(纽科姆)
non-metric methods(非量纲方法)
Old City(旧城)
One-mode data(1-模数据)
outdegree(点出度)
PAJEK,PAJEK(软件)
Pareto, V.(帕累托)
Parsons, T.(帕森斯)
partial networks(局部网)
path:in field theory(场论中的途径)
peak(顶点)
peripherality(边缘性)
plex(丛)
point centrality(点的中心度)
point file in GRADAP(GRADAP 中的点文件)
points in graphs(图论中的点)
positional analysis of networks(网络的位置分析)
positional approach,to data selection(数据选择的位置研究)
principle component analysis(PCA)(主成分分析)
projection in multidimensional scaling(多维量表中的投影)
proximity data(相近性数据)
Q-analysis(Q-分析)
Radcliffe-Brown, A.(拉德克利夫·布朗)
radius(半径)
random sub-graph(随机子图)
Rapoport, A.(拉波波特)
reyular(规则的)
relational data(关系数据)
reputational approach, to data selection(数据选择的声望研究)
role(角色)
rotation(旋转)
Sailer, L. D.(赛勒)
sampling(抽样)
scaling(量表)
Schwartz, J.(施瓦兹)
Schwartz, M.(施瓦兹)
Seidman, S. B.(赛德曼)
semi-cycles(半环)
separation(分离)
Shepard diagram(谢泼德图)
signed data and signed graphs(符号数据和符号图)
similarities(相似性)
Simmel, G.(齐美尔)
single attribute cluster(单一属性聚类)
single linkage cluster(单一联结聚类)
size of graph(图的规模)
slicing of matrices(矩阵的切分)
smallest space analysis(最小空间分析)